

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

- Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
- Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 13:20:20 ; Search time 232 Seconds
(without alignments)
268.620 Million cell updates/sec

Title: US-10-751-235-1

Perfect score: 393

Sequence: 1 LQPYADGSANMEAKFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table:

BLGapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgm2_1/USPTO.spool/US10751235/runat 12012006 120916 2581/app_query.fasta 1.263
-DB=Published Applications NA.New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPECL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MAIRIX=blobum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10751235 @CGN 1 1 122 @runat 12012006 120916 2581
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FEALOP=6 -FEALPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgm2_6/ptodata/2/pubpna/US08 NEW PUB.seq.*
- 2: /cgm2_6/ptodata/2/pubpna/US06 NEW PUB.seq.*
- 3: /cgm2_6/ptodata/2/pubpna/US07 NEW PUB.seq.*
- 4: /cgm2_6/ptodata/2/pubpna/US09 NEW PUB.seq.*
- 5: /cgm2_6/ptodata/2/pubpna/US10 NEW PUB.seq.*
- 6: /cgm2_6/ptodata/2/pubpna/US11 NEW PUB.seq.*
- 7: /cgm2_6/ptodata/2/pubpna/US12 NEW PUB.seq.*
- 8: /cgm2_6/ptodata/2/pubpna/US13 NEW PUB.seq.*
- 9: /cgm2_6/ptodata/2/pubpna/US14 NEW PUB.seq.*
- 10: /cgm2_6/ptodata/2/pubpna/US60 NEW PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68.5	17.4	2118	US-11-136-527-3214	Sequence 3214, Ap
2	67.5	17.2	3198	US-10-750-185-26357	Sequence 26357, A
3	67.5	17.2	3198	US-10-750-623-26357	Sequence 26357, A
4	64.5	16.4	3917	US-10-750-185-31952	Sequence 31952, A
5	64.5	16.4	3917	US-10-750-623-31952	Sequence 31952, A
6	64	16.3	804	US-10-793-626-1611	Sequence 1611, Ap
7	64	16.3	2918	US-10-793-626-3932	Sequence 3932, Ap
8	64	16.3	3660	US-10-793-626-4116	Sequence 4116, Ap

C	9	62	15.8	76329	7	US-11-150-888-16	Sequence 16, Appl
	10	61.5	15.6	2091	7	US-11-052-554A-484	Sequence 484, App
	11	61.5	15.6	2091	7	US-11-052-554A-485	Sequence 485, App
C	12	61.5	15.6	191797	7	US-11-121-086-13	Sequence 13, Appl
	13	61	15.5	2858	7	US-11-128-061-473	Sequence 473, App
	14	61	15.5	11131	6	US-10-240-708-27	Sequence 27, Appl
	15	60.5	15.4	1494	6	US-10-793-628-1623	Sequence 1623, Ap
	16	60.5	15.4	2005	7	US-11-136-527-3439	Sequence 3439, Ap
C	17	60.5	15.4	3721	6	US-10-793-626-3966	Sequence 3966, App
	18	59	15.0	1547	7	US-11-000-463-171	Sequence 171, App
	19	59	15.0	1956	7	US-11-052-554A-495	Sequence 495, App
C	20	59	15.0	2782	7	US-11-128-061-952	Sequence 952, App
	21	58.5	14.9	1016	6	US-10-750-185-42840	Sequence 42840, A
	22	58.5	14.9	1016	6	US-10-750-623-42840	Sequence 42840, A
	23	58.5	14.9	3027	6	US-10-467-657-185	Sequence 185, App
	24	58.5	14.9	3027	6	US-10-467-657-4555	Sequence 4555, App
C	25	58.5	14.9	137000	6	US-10-515-538-11	Sequence 11, Appl
	26	58.5	14.9	158692	7	US-11-121-086-30	Sequence 30, Appl
	27	58.5	14.9	1125000	6	US-10-995-561-13286	Sequence 13286, A
C	28	58	14.8	1260	6	US-10-750-185-52875	Sequence 52875, A
	29	58	14.8	1260	6	US-10-750-623-52875	Sequence 52875, A
C	30	57.5	14.6	879	6	US-10-750-185-30691	Sequence 30691, A
C	31	57.5	14.6	879	6	US-10-750-623-30691	Sequence 30691, A
	32	57.5	14.6	1987	6	US-10-750-185-49778	Sequence 49778, A
	33	57.5	14.6	1987	6	US-10-750-623-49778	Sequence 49778, A
	34	57.5	14.6	3590	7	US-11-136-527-3160	Sequence 3160, Ap
C	35	57	14.5	201	6	US-10-995-561-42931	Sequence 42931, A
C	36	57	14.5	2616	6	US-10-995-561-467	Sequence 467, App
C	37	57	14.5	2787	6	US-10-995-561-469	Sequence 469, App
C	38	57	14.5	2878	6	US-10-995-561-468	Sequence 468, App
C	39	57	14.5	3223	6	US-10-750-185-46677	Sequence 46677, A
C	40	57	14.5	3223	6	US-10-750-623-46677	Sequence 46677, A
C	41	56.5	14.4	1347	7	US-11-055-822-909	Sequence 909, App
C	42	56.5	14.4	1347	7	US-10-750-185-43888	Sequence 43888, A
C	43	56.5	14.4	2233	6	US-10-750-623-43888	Sequence 43888, A
C	44	56.5	14.4	3138	6	US-10-750-185-45175	Sequence 45175, A
	45	56.5	14.4	3138	6	US-10-750-623-45175	Sequence 45175, A

ALIGNMENTS

RESULT 1

US-11-136-527-3214
; Sequence 3214, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3214
; LENGTH: 2118
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-3214

Alignment Scores:
Pred. No.: 11.8 Length: 2118
Score: 68.50 Matches: 20
Percent Similarity: 46.15% Conservative: 10
Best Local Similarity: 30.77% Mismatches: 30
Query Match: 17.43% Indels: 5
DB: 7 Gaps: 2

US-10-751-235-1 (1-77) x US-11-136-527-3214 (1-2118)

QY 5 AlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThrLeuAspVal 24

```
Db 769 TCGGAGGAGGACCCACAGCATGAAACATCTTTGGGGCTACAGCATGATGTG 828
      :::::||||| :::::|||||
Qy 25 lIeGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThr---AspSerProVal 43
      :::::||||| :::::|||||
Db 829 ATCAGCCACTCATTTGGAGTGAATGTTGATTCCTTAACACCCACAGGACCTTTT 888
      :::::||||| :::::|||||
Qy 44 lIeGluAlaValTyThrAlaLeuLys-----GluAlaGluLeuArgSerThr 59
      :::::||||| :::::|||||
Db 889 GTGGAAAAGTCAAGAAGCTCTTAAATTTGATATCTTTGATCCATTGTTCTCTCAGTG 948
      :::::||||| :::::|||||
Qy 60 AspLeuLeuProTyThr 64
      :::::||||| :::::|||||
Db 949 ACACCTTTTCCATTC 963

RESULT 2
US-10-750-185-26357/c
; Sequence 26357, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 26357
; LENGTH: 3198
; TYPE: DNA
; ORGANISM: Bovine 19866881124994
US-10-750-185-26357

Alignment Scores:
Pred. No.: 30.3 Length: 3198
Score: 67.50 Matches: 17
Percent Similarity: 42.59% Conservative: 6
Best Local Similarity: 31.48% Mismatches: 28
Query Match: 17.18% Indels: 3
DB: 6 Gaps: 1

US-10-751-235-1 (1-77) x US-10-750-185-26357 (1-3198)
Qy 25 lIeGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThrAspSerProValIle 44
      :::::||||| :::::|||||
Db 2235 ATACGAGCTGAGTCTTTTAATATGTCGGGATAACGGGAACGTGACCTCTCTAGAGTG 2176
      :::::||||| :::::|||||
Qy 45 GluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeu----- 61
      :::::||||| :::::|||||
Db 2175 AGCCTTCAGTTGTGGCTCTCCATCACCCTCCCTCCCAACAACACTGAAACAGATT 2116
      :::::||||| :::::|||||
Qy 62 LeuProTyThrLysIleAspAlaLeuCysLysIleValPro 75
      :::::||||| :::::|||||
Db 2115 CTTCCTCTATGGGCTGTTCCTCAAGCTGCAAAATTATTTCCTCA 2074
      :::::||||| :::::|||||

RESULT 4
US-10-750-185-31952/c
; Sequence 31952, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31952
; LENGTH: 3917
; TYPE: DNA
; ORGANISM: Bovine 19866880527023
US-10-750-185-31952

Alignment Scores:
Pred. No.: 119 Length: 3917
Score: 64.50 Matches: 22
Percent Similarity: 45.68% Conservative: 15
Best Local Similarity: 27.16% Mismatches: 29
Query Match: 16.41% Indels: 15
DB: 6 Gaps: 4

US-10-751-235-1 (1-77) x US-10-750-185-31952 (1-3917)
```

```
Db 769 TCGGAGGAGGACCCACAGCATGAAACATCTTTGGGGCTACAGCATGATGTG 828
      :::::||||| :::::|||||
Qy 25 lIeGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThr---AspSerProVal 43
      :::::||||| :::::|||||
Db 829 ATCAGCCACTCATTTGGAGTGAATGTTGATTCCTTAACACCCACAGGACCTTTT 888
      :::::||||| :::::|||||
Qy 44 lIeGluAlaValTyThrAlaLeuLys-----GluAlaGluLeuArgSerThr 59
      :::::||||| :::::|||||
Db 889 GTGGAAAAGTCAAGAAGCTCTTAAATTTGATATCTTTGATCCATTGTTCTCTCAGTG 948
      :::::||||| :::::|||||
Qy 60 AspLeuLeuProTyThr 64
      :::::||||| :::::|||||
Db 949 ACACCTTTTCCATTC 963

RESULT 2
US-10-750-185-26357/c
; Sequence 26357, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 26357
; LENGTH: 3198
; TYPE: DNA
; ORGANISM: Bovine 19866881124994
US-10-750-185-26357

Alignment Scores:
Pred. No.: 30.3 Length: 3198
Score: 67.50 Matches: 17
Percent Similarity: 42.59% Conservative: 6
Best Local Similarity: 31.48% Mismatches: 28
Query Match: 17.18% Indels: 3
DB: 6 Gaps: 1

US-10-751-235-1 (1-77) x US-10-750-185-26357 (1-3198)
Qy 25 lIeGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThrAspSerProValIle 44
      :::::||||| :::::|||||
Db 2235 ATACGAGCTGAGTCTTTTAATATGTCGGGATAACGGGAACGTGACCTCTCTAGAGTG 2176
      :::::||||| :::::|||||
Qy 45 GluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeu----- 61
      :::::||||| :::::|||||
Db 2175 AGCCTTCAGTTGTGGCTCTCCATCACCCTCCCTCCCAACAACACTGAAACAGATT 2116
      :::::||||| :::::|||||
Qy 62 LeuProTyThrLysIleAspAlaLeuCysLysIleValPro 75
      :::::||||| :::::|||||
Db 2115 CTTCCTCTATGGGCTGTTCCTCAAGCTGCAAAATTATTTCCTCA 2074
      :::::||||| :::::|||||

RESULT 3
US-10-750-623-26357/c
; Sequence 26357, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
```



```

Qy      67 ile 67
Db      ::
        3047 CTT 3045

RESULT 6
US-10-793-626-1611
; Sequence 1611, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1611
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1611

```

Alignment Scores:	
Pred. No.:	14.8
Score:	64.00
Length:	804
Percent Similarity:	64.00
Matches:	19
Best Local Similarity:	44.62%
Conservative:	10
Mismatches:	18
Best Local Similarity:	29.23%
Mismatches:	18
Query Match:	16.28%
Indels:	18
DB:	6
Gaps:	3

	Qy	30	PheAsnTyrAnPheAepSerLeuThrThrAppSer-----ProValIle 44 :::
	Db	325	TTTAAATTATGTGCTAGATGGTGATGATATGGTAGACTTTAGATGTTTTGCGCCCGATTGA 384 :::
	Qy	45	GluAla-----valTYrrThralaleuLysGIuaIGlueArgserThr 59 : : :::

RESULT 7
US-10-793-626-3932/c
; Sequence 3932, Application US/10793626
; Publication No. US2005025478A1

```

;
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
;

```

RESULT 5
US-10-750-623-31952/c
; Sequence 31952, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:

; FILE REFERENCE: FM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623

US-10-750-623-31952	
; ORGANISM: Bovine	
19866880527023	
Alignment Scores:	
Pred. No.:	119
Score:	64.50
Percent Similarity:	45.68%
Best Local Similarity:	27.16%
Mismatches:	29
Conservative:	15
Matches:	22
Length:	3917

Db	3287	GAGAGTATTAACAGAGGGGGTTCGAAATAGGAAGACAGCAGCTTTTATGTCACCTCTAG	3228
Qy	22	LeuAspValIleGlyLeuSerLeuPheAsnTyAsnPheAspSerLeuThrThrAspSer	41
Db	3227	TTAAATATTATTACATTCACAAATTTTACACAAACATATGCACCATTTGAAGACAGAGGCT	3168
Qy	42	Pro-----ValIleGlu-----AlaValTyThrIaLeuTysGluAlaGlu	55
Db	3167	CCTGCCAGCATCTTCAGAAGGACCCATCCAGTGGTGGTCTCTCAGGTTAAGCTTTATCATC	3108

Qy	56	LeuArg	-----SerThrAspLeuLeuPro-----	TyrTrpLys	66
		:::			
Db	3107	CTTCAGGATGTTTTACTTCTACATAATCTCTCTCTGTAATAAACTTTGTATATGAAA			3048


```
; TYPE: DNA
; ORGANISM: Helicobacter pylori J99
US-11-052-554A-484

Alignment Scores:
Pred. No.: 142 Length: 2091
Score: 61.50 Matches: 22
Percent Similarity: 43.48% Conservatives: 8
Best Local Similarity: 31.88% Mismatches: 32
Query Match: 15.65% Indels: 7
DB: 7 Gaps: 1

US-10-751-235-1 (1-77) x US-11-052-554A-484 (1-2091)
QY 1 LeuGlnProTyAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 181 TTAACCAATACAAATTAATTAATTCCTTAGTCAATCTAGCCAGCGCTAGCGGATT 240
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyAsnPheAspSerLeuThrThrAsp 40
Db 241 ACCGGTGGATTGACAACTTAAGCTCAAGCGGATCAATCTCCTAGCGCTACCACTACT 300
QY 41 SerProValIleGluAlaValTyThrAlaLeuLysGluAlaLysGluAlaLysGluAlaLys 60
Db 301 TCTCGGCTATCAAGCTGTGGCTTAGCGCTCAATGCGCT----- 342
QY 61 LeuLeuProTyTrpLysIleAspAla 69
Db 343 ---GTGGGCATGTGGCAAGTCATAGCC 366

RESULT 11
US-11-052-554A-485
; Sequence 485, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 485
; LENGTH: 2091
; TYPE: DNA
; ORGANISM: Helicobacter pylori J99
US-11-052-554A-485

Alignment Scores:
Pred. No.: 142 Length: 2091
Score: 61.50 Matches: 22
Percent Similarity: 43.48% Conservatives: 8
Best Local Similarity: 31.88% Mismatches: 32
Query Match: 15.65% Indels: 7
DB: 7 Gaps: 1

US-10-751-235-1 (1-77) x US-11-052-554A-485 (1-2091)
QY 1 LeuGlnProTyAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 181 TTAACCAATACAAATTAATTAATTCCTTAGTCAATCTAGCCAGCGCTAGCGGATT 240
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyAsnPheAspSerLeuThrThrAsp 40
Db 241 ACCGGTGGATTGACAACTTAAGCTCAAGCGGATCAATCTCCTAGCGCTACCACTACT 300
QY 41 SerProValIleGluAlaValTyThrAlaLeuLysGluAlaLysGluAlaLysGluAlaLys 60
Db 301 TCTCGGCTATCAAGCTGTGGCTTAGCGCTCAATGCGCT----- 342
QY 61 LeuLeuProTyTrpLysIleAspAla 69
Db 343 ---GTGGGCATGTGGCAAGTCATAGCC 366

RESULT 12
US-11-121-086-13/c
; Sequence 13, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-13

Alignment Scores:
Pred. No.: 8,98e+04 Length: 191797
Score: 61.50 Matches: 23
Percent Similarity: 44.74% Conservatives: 11
Best Local Similarity: 30.26% Mismatches: 33
Query Match: 15.65% Indels: 9
DB: 7 Gaps: 3

US-10-751-235-1 (1-77) x US-11-121-086-13 (1-191797)
QY 2 GlnProTyAlaGluAspGlySerAlaValAsnMetGluAlaLys---PheSerGlnMet 20
Db 131256 CAACCTTATTAAGCAAGTCATGCTCTCTTTTCTAGCAAGCATCTATCTGCTT 131197
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyAsnPheAspSerLeuThrThrAsp 40
Db 131196 ACATTTCTGGCTATTGTGTCT-----TTCCACTTCAGTCTCTGCGAGATCA 131146
QY 41 SerProValIleGluAlaValTyThrAlaLeu-----LysGluAlaGlu 55
Db 131145 AGAAAGTCATAGCTTCAGCTTTTCCAGCTTTATTGTTTTCAGAGTGAAGTGTAT 131086
QY 56 LeuArgSerThrAspLeuLeuProTyTrpLysIleAspAlaLeuCys 71
Db 131085 GCTCTTCCAGCTTTCTAGGCTCCAAATGGAACCTGGAAGTCTATGT 131038

RESULT 13
US-11-128-061-473
; Sequence 473, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
```



```
Db      199 AATCTTGATGATCAAGAGTTTGGAAAGTGGTATCAAAATCAGACGGAATTTAATTTT 258
Qy      24  -----ValIleGly-----LeuSer 28
Db      259 AGTAATCCTCAATATTGGCTAGATTGTATTTGGTCATTATATGAAGTCATTATTACT 318
Qy      29  LeuPheAsnTyrAsnPheAspSerLeuThrThrAppSerProValIleGluAlaValTyr 48
Db      319 TATTATAATGACCAATTTTGATAATTTAACCTATTATCAATGATAAAGTACACAGAAATTTT 378
Qy      49  ThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeu 61
Db      379 ACA-----CAATCCGATGTTGATGACACAGATTTA 408
```

Search completed: January 15, 2006, 15:29:28
Job time : 249 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 05:07:45 ; Search time 474 Seconds
(without alignments)
1082.661 Million cell updates/sec

Title: US-10-751-235-1

Perfect score: 393

Sequence: 1 LQPVADGSVMEAKFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgm2_1/USPTO spool/US10751235/runat_12012006_120915_2508/app_query.fasta_1.263
-DB=N Geneseq -QFMT=fastap -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10751235 @CGN 1 1 727 @runat_12012006_120915_2508 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseq 21.*
- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	100.0	3377	13 ADT15447	Adt15447 Plant cDN
2	346	88.0	2467	14 AEB16909	Aeb16909 Thale cre
3	346	88.0	2467	14 AEB16911	Aeb16911 Thale cre
4	346	88.0	4170	14 AEB16910	Aeb16910 Thale cre

ALIGNMENTS

RESULT 1

ADT15447

ID ADT15447 standard; cDNA; 3377 BP.

XX AC ADT15447;

XX DT 13-JAN-2005 (first entry)

XX DE Plant cDNA, Seq ID 773.

XX KW Plant; ss; gene; transgenic; cold tolerance; growth rate;
KW drought tolerance; disease resistance; galactomannan production;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW lignin production; extreme osmotic condition tolerance;
KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
KW seed protein yield.

XX OS Viridiplantae.

XX PN US2004216190-A1.

XX PD 28-OCT-2004.

XX PF 18-DEC-2003; 2003US-00739930.

XX PR 28-APR-2003; 2003US-00424599.

XX PR 28-APR-2003; 2003US-00425115.

Aeb16928 Maize cyt
Adx61765 Plant ful
Aeb16927 Tomato cy
Aeb16926 wheat cyt
Aeb16925 Barley cy
Aeb16924 Rice cyto
Aeb16929 Sunflower
Aeb16950 Green alg
Aeb16948 Wheat cyt
Aeb16946 Barley cy
Aeb16945 Rice cyto
Aeb16944 Rice cyto
Aeb16943 Plant ful
Aeb16957 Soybean c
Aax60780 Soybean c
Aeb16943 Thale cre
Aeb16942 Thale cre
Aeb16941 Pinus rad
Aeb16940 Cinnamat
Aeb16939 Soybean c
Aeb16938 Thale cre
Aeb16937 Tomato cy
Aeb16936 Pea CYP97
Aeb16944 Rice cyto
Aeb16943 Diatom cy
Ady27732 Bacillus
Ady27460 Bacillus
Ady27356 Bacillus
Ady27458 Bacillus
Abk74941 Bacillus
Aas02937 Heme dona
Aad27354 Bacillus
Aas02940 Hybrid pr
Ady27456 Bacillus
Aaa46648 Nucleotid
Aad02365 Bacillus
Aaf54832 Nucleotid
Aaf26288 B. megate
Aaf26324 B. megate
Acc48000 B. megata
Abz70676 Bacillus
Aal54548 DNA of P4

PA (KOVA/) KOVALIC D K.
XX Kovalic DK;
XX WPI; 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
XX genetics, and in particular for producing transgenic plants with improved
XX biological characteristics.
XX
XX Claim 1; SEQ ID NO 773; 14pp; English.
XX
XX The invention relates a recombinant DNA construct comprising a
XX polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:
XX 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
XX (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,
XX Arabidopsis, wheat and rape but the specification does not indicate which
XX sequences is derived from which organism. Also included is a method of
XX producing a plant having an improved property, comprising transforming a
XX plant with a recombinant DNA construct comprising a promoter region
XX functional in a plant cell operably joined to a polynucleotide encoding a
XX polypeptide associated with the property, and growing the transformed
XX plant. The property is selected from improving plant cold tolerance, for
XX manipulating growth rate in plant cells by modification of the cell cycle
XX pathway, for improving plant drought tolerance, for providing increased
XX resistance to plant disease, for galactomannan production, for production
XX of plant growth regulators, for improving plant heat tolerance, for
XX improving plant tolerance to herbicides, for increasing the rate of
XX homologous recombination in plants, for lignin production, for improving
XX plant tolerance to extreme osmotic conditions, for improving plant
XX tolerance to pathogens or pests, for yield improvement by modification of
XX photosynthesis, for modifying seed oil yield and/or content, for
XX modifying seed protein yield and/or content, for yield improvement by
XX modification of carbohydrate, nitrogen or phosphorus use and/or uptake
XX and for yield improvement by providing improved plant growth and
XX development under at least one stress condition. The polynucleotide may
XX also encode a plant transcription factor. The methods and compositions of
XX the present invention are useful in the field of biochemistry and
XX genetics, in particular for producing transgenic plants with improved
XX biological characteristics such as increased yield, improved nitrogen
XX flow, increasing plant tolerance to cold or heat, improving plant
XX tolerance to extreme osmotic and drought conditions, and improving plant
XX tolerance to plant pests or pathogens. They can also be used in physical
XX arrays of molecules, plant breeding markers, computer-based storage and
XX analysis systems. The present sequence is one of the 5544 plant CDNA
XX sequences of the invention. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
XX Sequence 3377 BP; 967 A; 729 C; 736 G; 945 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,3e-49 Length: 3377
Score: 393.00 Matches: 77
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-751-235-1 (1-77) x APT15447 (1-3377)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 676 TTGCACCTTATGACAGACGCGAAGTGTGTGAATATGAGCGAAGTCTCTCAGATG 735
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
Db 736 ACACCTTGATGTCATTGGGTGTCTCTCTTTTAACTACAAATTCGATTCCTTGACTCTGAT 795
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 796 AGTCCTGTCTATTGAACTGTTTACATGCTCTCTTAAAGAGCTGAGCTTCGTTCTACTGAT 855

QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 856 CTTCTGCATATTGGAAGATCGATTCATTGTGTAAATAGTATGTCCTCCGACAG 906
RESULT 2
ID AEB16909 standard; cDNA; 2467 BP.
XX AEB16909;
XX
XX 08-SEP-2005 (first entry)
XX
XX Thale cress LUT1 (cytochrome P450 97C1) cDNA, SEQ ID NO: 5.
XX
XX Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
XX cytochrome P450 97C1; gene; ss; chromosome 3.
XX
XX Arabidopsis thaliana.
XX
XX Key Location/Qualifiers
XX CDS 1..2467
XX /*tag= a
XX /product= "Thale cress LUT1 protein"
XX /transl_except= (pos:496..650, aa:Thr-Ala)
XX /transl_except= (pos:942..1037, aa:Lys-Ile)
XX /transl_except= (pos:1242..1326, aa:Val-Ser)
XX /transl_except= (pos:1420..1523, aa:Lys-Asn)
XX /transl_except= (pos:1662..1841, aa:Val-Leu)
XX /transl_except= (pos:1839..2005, aa:Val-Trip)
XX /transl_except= (pos:2081..2183, aa:Lys-Phe)
XX /transl_except= (pos:2340..2440, aa:Asn-Gly)
XX
XX US2005150002-A1.
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIM/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16908.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
XX polypeptide having monooxygenase P450 activity, useful in altering the
XX carotenoid production in a plant for enhancing production of specific
XX carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 5; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
XX carotenoid monooxygenases in the cytochrome P450 family. The invention
XX also relates to altering carotenoid ratios in plants and microorganisms
XX using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
XX invention is useful in altering the carotenoid production in a plant for
XX enhancing production of specific carotenoid compounds that are potent
XX antioxidants. The present sequence is thale cress LUT1 [cytochrome P450
XX monooxygenase (CYP97C1); At3G53130 gene] cDNA. Note: The current sequence
XX is that of thale cress LUT1 cDNA which is located on chromosome 3.
XX
XX Sequence 2467 BP; 657 A; 491 C; 521 G; 798 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.79e-42 Length: 2467
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservatives: 0

Best Local Similarity: 71.03% Mismatches: 1
 Query Match: 88.04% Indels: 31
 DB: 14 Gaps: 1

US-10-751-235-1 (1-77) x AEB16909 (1-2467)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAenMetGluAlaLysPheSerGlnMet 20
 DB 747 TTGCAGCCTTATGCAAGACGGAAGTGTCTGTAATATGGAAGCGAAGTTCTCTCAGATG 806
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 DB 807 ACACCTGATGTCATGGGTGTCTCTTTTAACTACATTTGATCTTTGACTACTGAT 866
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 DB 867 AGTCCTGTCTATGAAGCTGTTTACACTGCTCTTAAAGAAGCTGAGCTTCGTTCTACTGAT 926
 QY 61 LeuLeuProTyrTpolysile----- 67
 DB 927 CTCTGCCCATATGGAAGGC-AAGTTTCTGTGTTTCTGTGTTTCTGTGTTTCTGTGTTG 985
 QY 68 -----AspAlaLeu 70
 DB 986 GAACAATTGGATCTTGTTAATTGAGAGGGTTTGGTTGTTTTTTTTCAGATCGATGCAATG 1045
 QY 71 CysLysIleValProArgGln 77
 DB 1046 TGTAAGATAGTCCCGAGACAG 1066

RESULT 3

AEB16911
 ID AEB16911 standard; cDNA; 2467 BP.

AC AEB16911;

DT 08-SEP-2005 (first entry)

DE Thale cress mutant LUT1 (cytochrome P450 97C1) cDNA, SEQ ID NO: 7.

KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
 KW cytochrome P450 97C1; ss; mutant.

OS Arabidopsis thaliana.

OS Synthetic.

OS US2005150002-A1.

OS 07-JUL-2005.

PF 02-JAN-2004; 2004US-00751235.

PR 02-JAN-2004; 2004US-00751235.

PA (DELL/) DELLAPENNA D.

PA (TIAN/) TIAN L.

PA (KIMJ/) KIM J.

PI Dellapenna D, Tian L, Kim J;

XX WPI; 2005-487984/49.

XX New expression vector comprising a nucleic acid sequence encoding a
 PT polypeptide having monooxygenase P450 activity, useful in altering the
 PT carotenoid production in a plant for enhancing production of specific
 PT carotenoid compounds.

XX Claim 10; SEQ ID NO 7; 135pp; English.

XX The present invention relates to genes, proteins and methods comprising
 CC carotenoid monooxygenases in the cytochrome P450 family. The invention
 CC also relates to altering carotenoid ratios in plants and microorganisms
 CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The

CC invention is useful in altering the carotenoid production in a plant for
 CC enhancing production of specific carotenoid compounds that are potent
 CC antioxidants. The present sequence is thale cress LUT1 [cytochrome P450
 CC monooxygenase (CYP97C1)] mutant cDNA.

SQ Sequence 2467 BP; 658 A; 491 C; 520 G; 798 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.: 7,798-42 Length: 2467
 Score: 346.00 Matches: 76
 Percent Similarity: 71.03% Conservative: 0
 Best Local Similarity: 71.03% Mismatches: 1
 Query Match: 88.04% Indels: 31
 DB: 14 Gaps: 1

US-10-751-235-1 (1-77) x AEB16911 (1-2467)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAenMetGluAlaLysPheSerGlnMet 20
 DB 747 TTGCAGCCTTATGCAAGACGGAAGTGTCTGTAATATGGAAGCGAAGTTCTCTCAGATG 806
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 DB 807 ACACCTGATGTCATGGGTGTCTCTTTTAACTACATTTGATCTTTGACTACTGAT 866
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 DB 867 AGTCCTGTCTATGAAGCTGTTTACACTGCTCTTAAAGAAGCTGAGCTTCGTTCTACTGAT 926
 QY 61 LeuLeuProTyrTpolysile----- 67
 DB 927 CTCTGCCCATATGGAAGGC-AAGTTTCTGTGTTTCTGTGTTTCTGTGTTTCTGTGTTG 985
 QY 68 -----AspAlaLeu 70
 DB 986 GAACAATTGGATCTTGTTAATTGAGAGGGTTTGGTTGTTTTTTTTCAGATCGATGCAATG 1045

RESULT 4

AEB16910

ID AEB16910 standard; DNA; 4170 BP.

AC AEB16910;

DT 08-SEP-2005 (first entry)

DE Thale cress LUT1 (cytochrome P450 97C1) DNA, SEQ ID NO: 6.

KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
 KW cytochrome P450 97C1; gene; ds; chromosome 3.

OS Arabidopsis thaliana.

OS US2005150002-A1.

OS 07-JUL-2005.

PF 02-JAN-2004; 2004US-00751235.

PR 02-JAN-2004; 2004US-00751235.

PA (DELL/) DELLAPENNA D.

PA (TIAN/) TIAN L.

PA (KIMJ/) KIM J.

PI Dellapenna D, Tian L, Kim J;

XX WPI; 2005-487984/49.

XX New expression vector comprising a nucleic acid sequence encoding a

XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX Unidentified.
XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 32608; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 1804 BP; 464 A; 392 C; 472 G; 476 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 1,44e-41 Length: 1804
Score: 343.00 Matches: 65
Percent Similarity: 92.21% Conservative: 6
Best Local Similarity: 84.42% Mismatches: 6
Query Match: 87.28% Indels: 0
DB: 13 Gaps: 0

US-10-751-235-1 (1-77) x ADX61765 (1-1804)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
|||||
DB 494 CTTGAGCCATATGCTTTGTGATGGGAAACCTGTCAATATGGAAGCGAGGTTTCTCAGTTG 553

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
|||||
DB 554 ACATTGGATGTGATTGTTTATCATATTGTTCACTACAAATTTTGTATTCCTCCACACAGAT 613
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
|||||
DB 614 AGTCCTGTGATGATGCTGTTTATATCTGCACCTCAAGAAGCAGAGCTTCTTCTACAGAT 673
QY 61 LeuLeuProTyrTriPlysIleAspAlaLeuCysLysIleValProArgGln 77
|||||
DB 674 CTTTGGCATACCTGGAAGGTGTTTCTTGTGCAAGATAATCCCAAGACAG 724

RESULT 7
AEB16927
ID AEB16927 standard; cDNA; 839 BP.
XX
AC AEB16927;
XX
DT 08-SEP-2005 (first entry)
XX
DE Tomato cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 25.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
XX
OS Lycopersicon esculentum.

FH Key Location/Qualifiers
CDS 1..837
FT /*tag= a
FT /product= "Tomato cytochrome P450 97C (CYP97C) protein"
FT /partial
FT /note= "No start and stop codons"
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.

XX 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.

PI Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
DR P-P8DB; AEB16921.
DR GENE BANK; BG643819.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 25; 135pp; English.

XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is tomato cytochrome P450
CC monooxygenase (CYP97C) cDNA.
XX

SQ Sequence 839 BP; 251 A; 160 C; 190 G; 238 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.04e-41 Length: 839

Score: 337.00 Matches: 66
Percent Similarity: 92.21% Conservative: 5
Best Local Similarity: 85.71% Mismatches: 6
Query Match: 85.75% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16927 (1-839)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 31 CTTTACCTGATGCAATTTCTGGCTCTGCAGTGAATATGAGGCAAAAGTTTCTCAACTA 90

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
Db 91 ACACCTGATGTTATTGGCTTGCACCTCTCAATTAATTTTGGATTCCTTACTACTGAC 150

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 151 AGTCAGTTATTGATGAGCTTTACACTGCCTAAAGAGCAGAACTCCGTTCAACTGAT 210

QY 61 LeuLeuProTyrTriPlysIleAspAlaLeuCysLysIleValProArgGln 77
Db 211 TTGTTGCCATATTGGCAGATCAAAGCTTTATGTAAGTTTCATCCACGACAA 261

RESULT 8
AEB16926
ID AEB16926 standard; cDNA; 1086 BP.
AC AEB16926;
XT 08-SEP-2005 (first entry)
DE Wheat cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 24.
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
OS Triticum aestivum.
FH Key Location/Qualifiers
FT CDS 1..1086
FT /tag= a
FT /product= "wheat cytochrome P450 97C (CYP97C) protein"
FT /partial
FT /note= "No stop codon"
XX US2005150002-A1.
XX 07-JUL-2005.
XX 02-JAN-2004; 2004US-00751235.
XX 02-JAN-2004; 2004US-00751235.
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIMJ/) KIM J.
XX Dellapenna D, Tian L, Kim J;
XX WPI; 2005-487984/49.
DR P-PSDB; AEB16920.
DR GENBANK; CA497665, BG906289, CA742365, CA742792.
XX New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX Claim 10; SEQ ID NO 24; 135pp; English.
XX The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention

CC also relates to altering carotenoid ratios in plants and microorganisms
CC using IUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is wheat cytochrome P450 monooxygenase
CC (CYP97C) cDNA.

XX Sequence 1086 BP; 243 A; 298 C; 295 G; 250 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.17e-40 Length: 1086
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16926 (1-1086)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 589 CTCGAGACTTATGCTTTGAGTGGTGAACCTGTTAATATGGAAGCGAGGTTTCTCAAAATG 648

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
Db 649 ACATTAGATGATGTTGGTTTATCCTTCAACTACAACTTTGATTCCTCATCATCATGAT 708

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 709 AGTCCTGTTATTGATGCTGTTTACACTGCACCTCAAGAAGCTGAGGCTGTTCTACAGAT 768

QY 61 LeuLeuProTyrTriPlysIleAspAlaLeuCysLysIleValProArgGln 77
Db 769 CTTTACCATTCTGCAGATCGATTTGCTGTGCAAGATTGTTCTTAGACAG 819

RESULT 9

AEB16925
ID AEB16925 standard; cDNA; 1638 BP.
XX AC AEB16925;
XX 08-SEP-2005 (first entry)
XX Barley cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 23.
XX Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
XX cytochrome P450 97C; gene; ss.
XX Hordeum vulgare; subsp. vulgare.
XX Hordeum vulgare; subsp. spontaneum.
XX Key Location/Qualifiers
XX CDS 1..1638
XX /tag= a
XX /product= "Barley cytochrome P450 97C (CYP97C) protein"
XX /transl_except= (pos:1585..1587, aa:Xaa)
XX /note= "Xaa may be any naturally occurring amino acid"
XX US2005150002-A1.
XX 07-JUL-2005.
XX 02-JAN-2004; 2004US-00751235.
XX 02-JAN-2004; 2004US-00751235.
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIMJ/) KIM J.
XX Dellapenna D, Tian L, Kim J;
XX WPI; 2005-487984/49.

```
DR P-PSDB; AEB16919.
DR GENBANK; BM816653, BU987393, CA023004.
DR DDBJ; AV835803.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 23; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using luti epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is barley cytochrome P450
CC monooxygenase (CYP97C) cDNA.
XX
SQ Sequence 1638 BP; 406 A; 403 C; 427 G; 401 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2,08e-40 Length: 1638
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservatives: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16925 (1-1638)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 574 CTGAGACATATGCTTTGAGCGGTGAACCTGTTAATATGGAAGCGAGATTTCTCAATG 633
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 634 ACACTAGATGTGATGTTGGTTGCTTTGTTCAACTCAACTGATTTGATTCCTCATCAGAT 693
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 694 AGTCCTGTTATGATGCTGTTTACACCGCAGCTGGAAGAGCAGAGGCTGTTCTACAGAT 753
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 754 CTTTACCACATACGCGCAGATGATTGCTGTCGCAAGATTGTTCTCAGACAG 804

RESULT 10
AEB16924
ID AEB16924 standard; cDNA; 1686 BP.
XX
AC AEB16924;
XX
DT 08-SEP-2005 (first entry)
XX
DE Rice cytochrome P450 97C2 (CYP97C2) cDNA, SEQ ID NO: 22.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C2;
KW cytochrome P450 97C2; gene; ss.
XX
OS Oryza sativa; japonica cultivar-group.
XX
FH Key Location/Qualifiers
FT CDS
FT 1..1686
FT /tag= a
FT /product= "Rice cytochrome P450 97C2 (CYP97C2) protein"
XX
US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.

XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
DR P-PSDB; AEB16918.
DR DDBJ; AK065689.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 22; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using luti epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is rice cytochrome P450 monooxygenase
CC (CYP97C2) cDNA.
XX
SQ Sequence 1686 BP; 393 A; 415 C; 459 G; 419 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,17e-40 Length: 1686
Score: 335.00 Matches: 66
Percent Similarity: 90.91% Conservatives: 4
Best Local Similarity: 85.71% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16924 (1-1686)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 622 CTTGAGACATCTGCTTTAAGTGGCAACCTGTAATATGGAAGCAAGGTTCTCTCAATG 681
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 682 ACTTAGATGTGATGTTGTTGCTTGTTCATTACAAATTTGATTCCTCATCAGAT 741
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 742 AGCCCTGTTATGATGCTGTTTACACTGCACCTCAAGGAAGCAGAACTTCGTTCTACAGAT 801
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 802 CTTTACCACATACGGAAGATTGATTGCTGTCGAAGATTGTTCTTAGACAA 852

RESULT 11
AEB16929
ID AEB16929 standard; cDNA; 624 BP.
XX
AC AEB16929;
XX
DT 08-SEP-2005 (first entry)
XX
DE Sunflower cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 27.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
XX
OS Helianthus annuus.
XX
FH Key Location/Qualifiers
```

```
FT CDS 1..624
FT /*tag= a
FT /product= "Sunflower cytochrome P450 97C (CYP97C)
FT /protein=
FT /partial
FT /note= "No start and stop codons"
XX
XX US2005150002-A1.
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16923.
XX GENBANK, BQ971938.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
XX polypeptide having monooxygenase P450 activity, useful in altering the
XX carotenoid production in a plant for enhancing production of specific
XX carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 27; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
XX carotenoid monooxygenases in the cytochrome P450 family. The invention
XX also relates to altering carotenoid ratios in plants and microorganisms
XX using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
XX invention is useful in affecting the carotenoid production in a plant for
XX enhancing production of specific carotenoid compounds that are potent
XX antioxidants. The present sequence is sunflower cytochrome P450
XX monooxygenase (CYP97C) cDNA.
XX
XX Sequence 624 BP; 186 A; 110 C; 144 G; 184 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.48e-38 Length: 624
XX Score: 319.00 Matches: 60
XX Percent Similarity: 89.61% Conservative: 9
XX Best Local Similarity: 77.92% Mismatches: 8
XX Query Match: 81.17% Indels: 0
XX DB: 14 Gaps: 0
XX
XX US-10-751-235-1 (1-77) x AEB16929 (1-624)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 253 CTAAGATCATACCGACGAGTGCACGCTCTGTTAATGAGCAACAGTTTCGCGAGTTA 312
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 313 ACCCTTGATGTTATCGGTCTAGCGGTATTAACTACAAATTTTCACTTACGCGCGAT 372
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 373 AGTCCTGTAATTGAATCTGTTTATACGCACTAAAGAAGCTGAAGCCGCTTCAACTGAT 432
QY 61 LeuLeuProTyrTrpIleAspAlaLeuCysLysIleValProArgGln 77
Db 433 CTTTGGCATATGGAGAGTAAGTGGCTTATGTAAGATTATACCAACAA 483
XX
XX RESULT 12
XX AEB16950
XX ID AEB16950 standard; cDNA; 1101 BP.
XX
```

```
AC AEB16950;
XX
XX 08-SEP-2005 (first entry)
XX
XX Green algae cytochrome P450 97A (CYP97A) like cDNA, SEQ ID NO: 48.
XX
XX DE Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
XX KW cytochrome P450 97A; gene; ss.
XX
XX OS Chlamydomonas reinhardtii.
XX
XX Key Location/Qualifiers
XX CDS 1..1101
XX FT /*tag= a
XX FT /product= "Green algae cytochrome P450 97A (CYP97A) like
XX FT protein"
XX FT /partial
XX FT /note= "No start and stop codons"
XX
XX US2005150002-A1.
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
XX (TIAN/) TIAN L.
XX (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16941.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
XX polypeptide having monooxygenase P450 activity, useful in altering the
XX carotenoid production in a plant for enhancing production of specific
XX carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 48; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
XX carotenoid monooxygenases in the cytochrome P450 family. The invention
XX also relates to altering carotenoid ratios in plants and microorganisms
XX using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
XX invention is useful in affecting the carotenoid production in a plant for
XX enhancing production of specific carotenoid compounds that are potent
XX antioxidants. The present sequence is green algae cytochrome P450
XX monooxygenase (CYP97A) like cDNA.
XX
XX Sequence 1101 BP; 217 A; 353 C; 342 G; 189 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.35e-21 Length: 1101
XX Score: 210.00 Matches: 41
XX Percent Similarity: 72.73% Conservative: 15
XX Best Local Similarity: 53.25% Mismatches: 21
XX Query Match: 53.44% Indels: 0
XX DB: 14 Gaps: 0
XX
XX US-10-751-235-1 (1-77) x AEB16950 (1-1101)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 100 CTAGACAGTATGCGCCCTCAGGCACCGCTGACATGGAACACTTCTTCACCGCGCTG 159
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 160 GGTCTGACATCATCGCGAAGCCGGTTCACACTAGACTTCGACTCGTGGCGCAGCAG 219
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
```

```
Db 220 GACCCGCTATCAGCCGCTGTACACCTTCTGCGGAGCGAGCACCGCTCCACAGCG 279
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 280 CCCATCGCCTACTGGAAACATCCCGGCATCCAGTTTGTGTGCGCGGCAG 330

RESULT 13
AEB16948
ID AEB16948 standard; cDNA; 1031 BP.
XX
AC AEB16948;
XX
DT 08-SEP-2005 (first entry)
XX
DE Wheat cytochrome P450 97A (CYP97A) cDNA, SEQ ID NO: 46.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
XX cytochrome P450 97A; gene; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 1..1031
FT /tag= a
FT /product= "Wheat cytochrome P450 97A (CYP97A) protein"
FT /transl_except= (pos:592..595, aa:Gln)
FT /partial
FT /note= "No start codon"
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
DR P-PSDB; AEB16939.
DR DDBJ; BJ234910.
DR GENEANK; CA736787, CA736801.
DR DDBJ; BJ238659, BJ233019.
DR GENEANK; CD882035.
XX
New expression vector comprising a nucleic acid sequence encoding a
polyptide having monooxygenase P450 activity, useful in altering the
carotenoid production in a plant for enhancing production of specific
carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 46; 135pp; English.
XX
The present invention relates to genes, proteins and methods comprising
carotenoid monooxygenases in the cytochrome P450 family. The invention
also relates to altering carotenoid ratios in plants and microorganisms
using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
invention is useful in altering the carotenoid production in a plant for
enhancing production of specific carotenoid compounds that are potent
antioxidants. The present sequence is wheat cytochrome P450 monooxygenase
(CYP97A) cDNA.
XX
SQ Sequence 1031 BP; 253 A; 237 C; 284 G; 257 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,52e-21 Length: 1031
Score: 207.00 Matches: 41
Percent Similarity: 71.43% Conservative: 14
```

```
Best Local Similarity: 53.25% Mismatches: 22
Query Match: 52.67% Indels: 0
DB: 14 Gaps: 0
US-10-751-235-1 (1-77) x AEB16948 (1-1031)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 364 CTGGACAAAGCGGCATCCGATGGGAGGATGTGGAGATGGAAATCTCTCTCTCGACTA 423
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 424 ACGCTGGATGTCATCCGGAAGCAGTGTTCATTAATGATTTTGATTCATTATCTTACGAT 483
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 484 AATGGAATAGTTGAGGCTGTGTATGTAACATTTACGGGAAGCGGAAATGCGGAGCACATCT 543
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 544 CTTATTCCACTTGGGAAATATCCCATATGGAAGACATCTCCCTCGGCAG 594

RESULT 14
AEB16946
ID AEB16946 standard; cDNA; 1527 BP.
XX
AC AEB16946;
XX
DT 08-SEP-2005 (first entry)
XX
DE Barley cytochrome P450 97A (CYP97A) cDNA, SEQ ID NO: 44.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
XX cytochrome P450 97A; gene; ss.
XX
OS Hordeum vulgare; subsp. vulgare.
XX OS Hordeum vulgare; subsp. spontaneum.
XX
FH Key Location/Qualifiers
FT CDS 1..1527
FT /tag= a
FT /product= "Barley cytochrome P450 97A (CYP97A) protein"
FT /partial
FT /note= "No start codon"
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
DR P-PSDB; AEB16937.
DR DDBJ; AV939715, AV941342, AV939552, AV939556.
DR GENEANK; CA004011.
DR DDBJ; BJ480615, BJ485000, BJ448041, BJ455787, AV910152, AV938407.
DR EMBL; AJ477620, AJ477618, AJ477619, AJ832622.
XX
New expression vector comprising a nucleic acid sequence encoding a
polyptide having monooxygenase P450 activity, useful in altering the
carotenoid production in a plant for enhancing production of specific
carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 44; 135pp; English.
XX
```

CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is barley cytochrome P450
CC monooxygenase (CYP97A) cDNA.

XX
XX
SQ Sequence 1527 BP; 399 A; 343 C; 395 G; 390 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.11e-21 Length: 1527
Score: 207.00 Matches: 41
Percent Similarity: 71.43% Conservative: 14
Best Local Similarity: 53.25% Mismatches: 22
Query Match: 52.67% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16946 (1-1527)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
DB 361 CTCGACAAAGGCTGCTTCGACGGGAGGATGGAGATGGAAATCTCTTCTCCGACTA 420
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 421 ACGCTGGATGTCATCGGGAAGCGGTGTTCAATTATGATTTTGAATTCATTATCTTACGAT 480
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 481 AATGGAATAGTTGAGCTGTGTATGTAACACTGCGGGAAGCAATGCGGAGTACATCT 540
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
DB 541 CCTATTCCAACATGGGAATACCCATATGGAAAGACATCTCCCTCGGCAG 591

RESULT 15
AEB16945
ID AEB16945 standard; DNA; 1899 BP.
XX
XX AEB16945;
AC
DT 08-SEP-2005 (first entry)
XX
DE Rice cytochrome P450 97A (CYP97A) DNA, SEQ ID NO: 43.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
KW cytochrome P450 97A; gene; ds.
XX
OS Oryza sativa; japonica cultivar-group.
XX
FH Key Location/Qualifiers
FT CDS 1..1899
FT /*tag= a
FT /product= "Rice cytochrome P450 97A (CYP97A) protein"
XX
FN US2005150002-A1.
XX
XX
PD 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
PI
XX
XX WPI; 2005-487984/49.
DR
DR P-PSDB; AEB16936.

DR DDBJ; AP004028.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 43; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is rice cytochrome P450 monooxygenase
CC (CYP97A) DNA.

XX
XX
SQ Sequence 1899 BP; 454 A; 470 C; 504 G; 471 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.18e-20 Length: 1899
Score: 206.00 Matches: 41
Percent Similarity: 71.43% Conservative: 14
Best Local Similarity: 53.25% Mismatches: 22
Query Match: 52.42% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AEB16945 (1-1899)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
DB 733 TTGGACAAGGCGACACCGATGGGAGGATGGAGATGGAAATCTTCTCTCGACTA 792
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 793 ACACCTGGATGTCATTGGGAAGCGATCTTCAATTATGATTTTCACTCATTTGCTTACGAT 852
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 853 AATGGAATAGTTGAGGCGAGTGTATGTGACACTGCGGGAAGCAATGCGGAGCACTTCT 912
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
DB 913 CCTATACCAACTTGGGAATACCCATATGGAAAGATATTTCCTCCGCGGCAG 963

Search completed: January 15, 2006, 13:20:16
Job time : 477 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 12:59:30 ; Search time 171 Seconds
(without alignments)
800.423 Million cell updates/sec

Title: US-10-751-235-1
Perfect score: 393
Sequence: 1 LQVABDSAVNMEAFKFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235 @CGN 1 1 193 @runat 12012006 120916 2538 -NCPU=6 -ICPU=3
-NO.MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgm2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgm2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgm2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgm2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgm2_6/ptodata/1/ina/H COMB.seq.*
6: /cgm2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgm2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgm2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	200	50.9	1831	3 US-08-948-564-15
2	196	49.9	404	3 US-09-615-192A-156
3	196	49.9	404	3 US-09-169-789-156
4	76	19.3	3150	3 US-10-018-730A-3
5	74	18.8	1633	3 US-09-583-447A-7
6	71.5	18.2	1349	3 US-09-583-447A-5
7	71.5	18.2	1515	3 US-09-583-447A-3
8	71.5	18.2	1659	3 US-09-583-447A-1
9	71.5	18.2	1973	3 US-09-583-447A-11

ALIGNMENTS

RESULT 1

US-08-948-564-15
; Sequence 15, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminszky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400

```

; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20..1747
US-08-948-564-15

Alignment Scores:
Pred. No.: 2,57e-20 Length: 1831
Score: 200.00 Matches: 40
Percent Similarity: 73.24% Conservativeness: 12
Best Local Similarity: 56.34% Mismatches: 19
Query Match: 50.89% Indels: 0
DB: 3 Gaps: 0

US-10-751-235-1 (1-77) x US-08-948-564-15 (1-1831)
QY 7 AspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIleGly 26
Db 647 GACTCAATTGAATGGATCTCTGAGGCGAGATTTCTAGTTGGCTCTGATATTATGGG 706

QY 27 LeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGluAla 46
Db 707 CTTGGTGTTCACTATGACCTTTGGTCTCTGTCACCAAGATCCAGTATTAAAGGCA 766

QY 47 ValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeuLeuProTyrTrpLys 66
Db 767 GTCTATGGCACTCTTTTGAAGCTGAACACAGATCCACTTCTTACATTCATATTGGAA 826

QY 67 IleAspAlaLeuCysLysIleValProArgGln 77
Db 827 ATTCCATTGGCAAGGTGGATAGTCCCAAGGCAA 859

RESULT 2
US-09-615-192A-156
; Sequence 156, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 156
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-156

Alignment Scores:
Pred. No.: 1.26e-20 Length: 404
Score: 196.00 Matches: 39
Percent Similarity: 74.63% Conservativeness: 11
Best Local Similarity: 58.21% Mismatches: 17
Query Match: 49.87% Indels: 0
DB: 3 Gaps: 0

US-10-751-235-1 (1-77) x US-09-615-192A-156 (1-404)
QY 11 ValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIleGlyLeuSerLeuPhe 30
Db 31 GTTGACATGGAGGAGCAATTTTCAAATCTAGCTTTGGACATATTGGATTGTGTATTT 90

QY 31 AsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGluAlaValTyrThrAla 50
Db 91 AACTATGATTTTGGATCCGTTACTCGAGAAATCAACAGTAATCAAGGCAGTCTATGGTACA 150

QY 51 LeuLysGluAlaGluLeuArgSerThrAspLeuLeuProTyrTrpLysIleAspAlaLeu 70
Db 151 TTGTTTGAAGCTGAGCATAGATCAACCTTTTACATACCATCTGGAATTTCCGCTGGCA 210

QY 71 CysLysIleValProArgGln 77
Db 211 AGATGGTTAGTTCTTCGCCAA 231

RESULT 4
US-10-018-730A-3
; Sequence 3, Application US/10018730A
; Patent No. 6794168
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C2
; CURRENT APPLICATION NUMBER: US/09/169,789
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 156
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-156

Alignment Scores:
Pred. No.: 1.26e-20 Length: 404
Score: 196.00 Matches: 39
Percent Similarity: 74.63% Conservativeness: 11
Best Local Similarity: 58.21% Mismatches: 17
Query Match: 49.87% Indels: 0
DB: 3 Gaps: 0

US-10-751-235-1 (1-77) x US-09-169-789-156 (1-404)
QY 11 ValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIleGlyLeuSerLeuPhe 30
Db 31 GTTGACATGGAGGAGCAATTTTCAAATCTAGCTTTGGACATATTGGATTGTGTATTT 90

QY 31 AsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGluAlaValTyrThrAla 50
Db 91 AACTATGATTTTGGATCCGTTACTCGAGAAATCAACAGTAATCAAGGCAGTCTATGGTACA 150

QY 51 LeuLysGluAlaGluLeuArgSerThrAspLeuLeuProTyrTrpLysIleAspAlaLeu 70
Db 151 TTGTTTGAAGCTGAGCATAGATCAACCTTTTACATACCATCTGGAATTTCCGCTGGCA 210

QY 71 CysLysIleValProArgGln 77
Db 211 AGATGGTTAGTTCTTCGCCAA 231

RESULT 4
US-10-018-730A-3
; Sequence 3, Application US/10018730A
; Patent No. 6794168
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C2
; CURRENT APPLICATION NUMBER: US/09/169,789
; PRIOR FILING DATE: 1998-10-09
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 156
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-156

Alignment Scores:
Pred. No.: 1.26e-20 Length: 404
Score: 196.00 Matches: 39
Percent Similarity: 74.63% Conservativeness: 11
Best Local Similarity: 58.21% Mismatches: 17
Query Match: 49.87% Indels: 0
DB: 3 Gaps: 0
```

```

: APPLICANT: Luet, Wong
: APPLICANT: Jonathan, Jones
: TITLE OF INVENTION: PROCESS FOR OXIDIZING AROMATIC COMPOUNDS
: FILE REFERENCE: P02353US1 / 10112404 / N.76277B
: CURRENT APPLICATION NUMBER: US/10/018,730A
: CURRENT FILING DATE: 2002-04-04
: PRIOR APPLICATION NUMBER: GB 9914373.7
: PRIOR FILING DATE: 1999-06-18
: PRIOR APPLICATION NUMBER: PCT/GB00/023379
: PRIOR FILING DATE: 2000-06-19
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatIn version 3.1
: SEQ ID NO 3
: LENGTH: 3150
: TYPE: DNA
: ORGANISM: Bacillus megaterium
: US-10-018-730A-3

```

Alignment Scores:	
Pred. No.:	0.432
Score:	76.00
Percent Similarity:	56.52%
Best Local Similarity:	39.13%
Query Match:	19.34%
DB:	3
Length:	3150
Matches:	18
Conservative:	8
Mismatches:	18
Indels:	2
Gaps:	1

US-10-751-235-1 (1-77) X US-10-018-730A-3 (1-3150)

[illegible]

RESULT 5

```

US-09-583-447A-7
; Sequence 7, Application US/09583447A
; Patent No. 6645745
; GENERAL INFORMATION:
; APPLICANT: WOJNOWSKI, Leszek
; APPLICANT: GELLNER, Klaus
; APPLICANT: EISELT, Regina
; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
; FILE OF INVENTION: (CYP3A) GENE FAMILY: CYP3AX
; FILE REFERENCE: 310115.401
; CURRENT APPLICATION NUMBER: US/09/583,447A
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 7
; LENGTH: 1633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(750)
US-09-583-447A-7

```

Alignment Scores:		
Pred. No.:	0.347	Length: 1633
Score:	74.00	Matches: 21
Percent Similarity:	50.00%	Conservative: 15
Best Local Similarity:	29.17%	Mismatches: 34
Query Match:	18.83%	Indels: 2
DB:	3	Gaps: 2

US-10-751-235-1 (1-77) x US-09-583-447A-7 (1-1633)

Qy	1	LeuGlnProTyrAlaGluAspGlySerAlaValAenMetGluAlaLysPheSerGlnMet	20
Db	478	CTGAGCAGGAGAGCAGACAGCAGCAAGTCCATCAACTGAAAGATTTCTTTGGGGCTAC	537
Qy	21	ThrLeuaspValIleGleLeuSerLeuPheAenTyrAenPheAspSerLeuThrThr---	39
Db	538	ACCATGGATGTAATCACTGGGCACATATTGGAGTGAACCTTGGATCTCTCAACAATCCA	597
Qy	40	AspSerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThr	59
Db	598	CAAGATCCCTTCTTGAAAAATATGAAGAAGCTTTTAAAA---TTGGATTTTTTGGATCCC	654
Qy	60	AspLeuLeuProTyrTrpLysIleAspAlaLeuCys	71
Db	655	TTTTTTACTCTTAATATACAGAGTCTCGGTGGTGTTCG	690

RESULT 6

```

US-09-583-447A-5
; Sequence 5, Application US/09583447A
; Patent NO. 6645745
; GENERAL INFORMATION:
; APPLICANT: WOJNOMSKI, Leszek
; APPLICANT: GELLNER, Klaus
; APPLICANT: EISELT, Regina
; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
; TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3AX
; FILE REFERENCE: 310115.401
; CURRENT APPLICATION NUMBER: US/09/583,447A
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1349
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1260)
; US-09-583-447A-5

```

Alignment Scores:	
Pred. No.:	0.639
Score:	71.50
Percent Similarity:	54.72%
Best Local Similarity:	33.96%
Query Match:	18.19%
DB:	3
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:
	1349
	18
	11
	23
	1
	1

US-10-751-235-1 (1-77) x US-09-583-447A-5 (1-1349)

Qy	1	LeuGlnProTyrAlaGluAspGlySerAlaValenMetGluAlaLysPheSerGlnMet	20
Db	478	CTGAGCGGAAGCAGAGAACAGCAAGTCACCTTGAAAGATTTCTTTGGGCCTAC	537
Qy	21	ThrLeuaspVallleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThr	39
Db	538	ACCATGGATGTAACTCACTGGCACATTATTGGAGTGAACCTTGGATCTCTCAACAATCCA	597
Qy	40	AspSerProVallleGluAlaValTyrThrAlaLeuLys	52
Db	598	CAAGATCCCTTCTTGAAGAATATCAAGAAGCGTTTTAAA	636

RESULT 7

```

US-09-583-447A-3
; Sequence 3, Application US/09583447A
;
; GENERAL INFORMATION:
; Patent No. 6645745
;
; APPLICANT: WOJNOWSKI, Leszek
; APPLICANT: GELLNER, Klaus
; APPLICANT: EISELT, Regina
; TITLE OF INVENTION: IDENTIFICATION
; TITLE OF INVENTION: (CYP3A) GENE F
;

```

FILE REFERENCE: 310115.401
CURRENT APPLICATION NUMBER: US/09/583,447A
CURRENT FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 1515
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1515)
US-09-583-447A-3
Alignment Scores:
Pred. No.: 0.751 Length: 1515
Score: 71.50 Matches: 18
Percent Similarity: 54.72% Conservative: 11
Best Local Similarity: 33.96% Mismatches: 23
Query Match: 18.19% Indels: 1
Gaps: 3
DB:
US-10-751-235-1 (1-77) x US-09-583-447A-3 (1-1515)
Qy 1 LeuGlnProTyAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 478 CTGAGCGAGGAGCAGACAGCAAGTCCATCACTTGAAGAGTTCTTTGGGGCCTAC 537
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThrThr--- 39
Db 538 ACCATGGATGTAATCACTGCGCACATTTATTTGGAGTGAATCTTCTCAACAATCCA 597
Qy 40 AspSerProValIleGluAlaValTyThrAlaLeuLys 52
Db 598 CAAGATCCCTTCTCGAAAAATATGAAGAAGCTTTTAAAA 636
RESULT 8
US-09-583-447A-1
Sequence 1, Application US/09583447A
Patent No. 6645745
GENERAL INFORMATION:
APPLICANT: WOJNOWSKI, Leszek
APPLICANT: GELLNER, Klaus
APPLICANT: EISELT, Regina
TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3AX
FILE REFERENCE: 310115.401
CURRENT APPLICATION NUMBER: US/09/583,447A
CURRENT FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1659
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (105)..(1616)
US-09-583-447A-1
Alignment Scores:
Pred. No.: 0.853 Length: 1659
Score: 71.50 Matches: 18
Percent Similarity: 54.72% Conservative: 11
Best Local Similarity: 33.96% Mismatches: 23
Query Match: 18.19% Indels: 1
Gaps: 3
DB:
US-10-751-235-1 (1-77) x US-09-583-447A-1 (1-1659)
Qy 1 LeuGlnProTyAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 582 CTGAGCGAGGAGCAGACAGCAAGTCCATCACTTGAAGAGTTCTTTGGGGCCTAC 641

Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThrThr--- 39
Db 642 ACCATGGATGTAATCACTGCGCACATTTATTTGGAGTGAATCTTCTCAACAATCCA 701
Qy 40 AspSerProValIleGluAlaValTyThrAlaLeuLys 52
Db 702 CAAGATCCCTTCTCGAAAAATATGAAGAAGCTTTTAAAA 740
RESULT 9
US-09-583-447A-11
Sequence 11, Application US/09583447A
Patent No. 6645745
GENERAL INFORMATION:
APPLICANT: WOJNOWSKI, Leszek
APPLICANT: GELLNER, Klaus
APPLICANT: EISELT, Regina
TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3AX
FILE REFERENCE: 310115.401
CURRENT APPLICATION NUMBER: US/09/583,447A
CURRENT FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 1973
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(222)
OTHER INFORMATION: n=a, c, t or g
US-09-583-447A-11
Alignment Scores:
Pred. No.: 1.09 Length: 1973
Score: 71.50 Matches: 18
Percent Similarity: 54.72% Conservative: 11
Best Local Similarity: 33.96% Mismatches: 23
Query Match: 18.19% Indels: 1
Gaps: 3
DB:
US-10-751-235-1 (1-77) x US-09-583-447A-11 (1-1973)
Qy 1 LeuGlnProTyAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 940 CTGAGCGAGGAGCAGACAGCAAGTCCATCACTTGAAGAGTTCTTTGGGGCCTAC 999
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThrThr--- 39
Db 1000 ACCATGGATGTAATCACTGCGCACATTTATTTGGAGTGAATCTTCTCAACAATCCA 1059
Qy 40 AspSerProValIleGluAlaValTyThrAlaLeuLys 52
Db 1060 CAAGATCCCTTCTCGAAAAATATGAAGAAGCTTTTAAAA 1098
RESULT 10
US-08-802-141-3
Sequence 3, Application US/08802141
Patent No. 5773009
GENERAL INFORMATION:
APPLICANT: GLASS, ROGER I.
APPLICANT: GENTSCH, JOHN R.
APPLICANT: BHAN, M. K.
APPLICANT: DAS, BIMAL K.
TITLE OF INVENTION: ROTAVIRUS STRAIN AND RELATED
TITLE OF INVENTION: COMPOSITIONS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEEDLE & ROSENBERG, P.C.
STREET: 127 Peachtree Street, Suite 1200
CITY: Atlanta
STATE: Georgia

;; COUNTRY: USA
;; ZIP: 30303
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/802,141
;; FILING DATE: 19-FEB-1997
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/231,041
;; FILING DATE: 15-APR-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Spratt, Gwendolyn D.
;; REGISTRATION NUMBER: 36,016
;; REFERENCE/DOCKET NUMBER: 1414.609
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 404/688-0770
;; TELEFAX: 688-9880
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2352 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 11..2329
US-08-802-141-3

Alignment Scores:
Pred. No.: 1.66 Length: 2352
Score: 71.00 Matches: 21
Percent Similarity: 46.77% Conservative: 8
Best Local Similarity: 33.87% Mismatches: 23
Query Match: 18.07% Indels: 10
DB: 2 Gaps: 3

US-10-751-235-1 (1-77) x US-08-802-141-3 (1-2352)

QY 4 TyralaGluAspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMetThrLeuAsp 23
Db 920 TATATGAGGCGATGCATAGACTT-----GTGCTCATACTACTGTTCA 964
QY 24 ValIleGlyLeuSerLeuPheAsnTyraAsnPheAspSerLeuThrThrAspSerProVal 43
Db 965 GTGAACGGTGTGTATGATGATATTAACACTGGGTCGTTACCAACTGATTCACGATC 1024
QY 44 IleGluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeuPro 63
Db 1025 CGAAAT-----TACACGTTTGTGAAGAAGTTCGTTGATATACGTTGAT----- 1069
QY 64 Tyrtip 65
Db 1070 TATTCG 1075

RESULT 11
US-08-911-321-6
; Sequence 6, Application US/08911321
; Patent No. 6010703
; GENERAL INFORMATION:
; APPLICANT: Roger K. Maes and Stephen J. Spatz
; TITLE OF INVENTION: Recombinant Forvirus
; TITLE OF INVENTION: Vaccine Against
; TITLE OF INVENTION: Feline Rhinotracheitis
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos

;; STATE: Michigan
;; COUNTRY: USA
;; ZIP: 48864
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
;; MEDIUM TYPE: storage
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/911,321
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/096,183
;; FILING DATE: July 26, 1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ian C. McLeod
;; REGISTRATION NUMBER: 20,931
;; REFERENCE/DOCKET NUMBER: MSU 4.1-166
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (517) 347-4100
;; TELEFAX: (517) 347-4103
;; TELEX: No. 6010703e
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6176
;; TYPE: Nucleic Acid
;; STRANDEDNESS: Double
;; TOPOLOGY: Linear
;; MOLECULE TYPE:
;; DESCRIPTION: DNA
;; HYPOTHEICAL: NO
;; ANTI-SENSE: No
;; ORIGINAL SOURCE:
;; ORGANISM: Feline herpesvirus-1
;; STRAIN: 1
;; INDIVIDUAL ISOLATE: C-27
;; CELL TYPE: N/A
;; FEATURE:
;; NAME/KEY:
;; LOCATION:
;; IDENTIFICATION METHOD: Sequencing
;; OTHER INFORMATION: DNA encoding PK, gG, gD,
;; OTHER INFORMATION: gI, gE
US-08-911-321-6
Alignment Scores:
Pred. No.: 6.41 Length: 6176
Score: 71.00 Matches: 19
Percent Similarity: 45.61% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 15
Query Match: 18.07% Indels: 16
DB: 3 Gaps: 2
US-10-751-235-1 (1-77) x US-08-911-321-6 (1-6176)
QY 3 ProTyralaGluAspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMetThr--- 21
Db 3668 CGGAAATGGAAGATGGTGAATCTACGCACTGGCGGTAAAGTTTAAACATAACAAA 3727
QY 22 LeuAspValIleGlyLeuSerLeuPheAsnTyraAsnPheAspSer----- 36
Db 3728 GCTGATGATTTGGCGCTTTCGTTGTTTACTCATTCGATACGCGTGCATCGACAT 3787
QY 37 -----LeuThrThrAspSerProVal 43
Db 3788 CATCGGCGAGAAATTTGAATGGTGAATTTCTTACTCTCCATCCCGATG 3838
RESULT 12
US-09-248-796A-4077
; Sequence 4077, Application US/09248796A
; Patent No. 6747137

```
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4077
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4077

Alignment Scores:
Pred. No.:          0.261          Length:          432
Score:              69.50          Matches:         18
Percent Similarity: 53.97%          Conservative:    16
Best Local Similarity: 28.57%          Mismatches:     24
Query Match:        3              Indels:          5
DB:                  3              Gaps:           2

US-10-751-235-1 (1-77) x US-09-248-796A-4077 (1-432)
QY 14 GluAlaLysPheSerGlnMetThrLeuAspValIleGlyLeuSerLeuPheAsn----- 31
Db 52 GATGCCAAATTAGTCAGGTTAATCAATAATCAAT-----TTATTTTCATGAAGAA 102
QY 32 TyrAsnPheAspSerLeuThrThrAspSerProValIleGluAlaValTyrThrAlaLeu 51
Db 103 TATACCATTTACTGATATACCTAAAGACAGACAAATTTTACAGTACTCGATTTCTAGCAAC 162
QY 52 LysGluAlaGluLeuArgSerThrAspLeuProTyrTyrIleAspAlaLeuCys 71
Db 163 AAAAAGGGCTCAATAAAGTTTCAGAGAGATATTGCCAAATCAATAATGACTTGGCGTGT 222
QY 72 LysIleVal 74
Db 223 ACCATTTA 231

RESULT 13
US-09-712-363-92
; Sequence 92, Application US/09712363
; Patent No. 6892139
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
```

```
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-92

Alignment Scores:
Pred. No.:          3.08          Length:          2217
Score:              69.00          Matches:         21
Percent Similarity: 47.83%          Conservative:    12
Best Local Similarity: 30.43%          Mismatches:     25
Query Match:        3              Indels:         11
DB:                  3              Gaps:           2

US-10-751-235-1 (1-77) x US-09-712-363-92 (1-2217)
QY 6 GluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIle 25
Db 301 GACGATGGCGAGCGCTTG-----CGGTCCTTCTGAGTCAGCTGTCG----- 345
QY 26 GlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGlu 45
Db 346 -----GGCTACTCCATCGAGGACGTACTACTCGAGGACTCTCTGCTCGAT 390
QY 46 AlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeuPro-TyrTr 65
Db 391 GCGGTCGAGGAGCGCTGCTGACCTGTCACCGAATCGCGGACCTGTACTCTGCTGAC 450
QY 65 LysIleAspAlaLeuCysValIle 73
Db 451 GGGTCGGGGCTCCGCTCGGGCATC 475

RESULT 14
US-07-603-133B-19
; Sequence 19, Application US/07603133B
; Patent No. 5298244
; GENERAL INFORMATION:
; APPLICANT: Redmond, Mark J.
; APPLICANT: Ijaz, Mohammed K.
; APPLICANT: Parker, Michael D.
; TITLE OF INVENTION: ASSEMBLED VIRAL PARTICLES AND THEIR
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/603,133B
; FILING DATE: 19901025
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9313-0004.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-7250
; TELEFAX: (415) 327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 2363 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 10..2337
US-07-603-133B-19

Alignment Scores:
Pred. No.: 3.37 Length: 2363
Score: 69.00 Matches: 21
Percent Similarity: 45.16% Conservatives: 7
Best Local Similarity: 33.87% Mismatches: 24
Query Match: 17.56% Indels: 10
DB: 2 Gaps: 3

US-10-751-235-1 (1-77) x US-07-603-133B-19 (1-2363)

QY 4 TyrAlaGluAspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMetThrLeuAsp 23
Db 922 TATACAGAGATGGTGAAGAAGTTACT-----GCACATACACTACGTGTTC A 966
QY 24 ValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAspSerProVal 43
Db 967 GTAATGGAAATGAATGATTTTAATTATATATATGGTGATCATTAACCGACTGAT-----TTC 1020
QY 44 IleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeuLeuPro 63
Db 1021 GTAATATCAATATGAAGTGAATTAGGAA-----AATCTTTGTGTATATAGAC 1071
QY 64 TyrTIP 65
Db 1072 TACTG3 1077

RESULT 15

US-08-653-740-2
Sequence 2, Application US/08653740
Patent No. 5792850
GENERAL INFORMATION:
APPLICANT: James W. Baumgartner
APPLICANT: Donald C. Foster
APPLICANT: Frank J. Grant
APPLICANT: Cindy A. Sprecher
TITLE OF INVENTION: HEMATOPOIETIC CYTOKINE RECEPTOR
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/653,740
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31,648
REFERENCE/DOCKET NUMBER: 95-31
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6673
TELEFAX: 206-442-6678
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2368 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 23..1759
US-08-653-740-2

Alignment Scores:
Pred. No.: 3.38 Length: 2368
Score: 69.00 Matches: 24
Percent Similarity: 47.37% Conservatives: 12
Best Local Similarity: 31.58% Mismatches: 33
Query Match: 17.56% Indels: 7
DB: 2 Gaps: 2
US-10-751-235-1 (1-77) x US-08-653-740-2 (1-2368)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 1940 CTTGGCCCCCATGGGGGAAGACACACGATGGAGGTGGAGCAAGAGAAATACATGAATTTG 1999
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsn-----Phe-AspSerLe 37
Db 2000 AGAGTGGCAGCTGCTGCCAAATCTGTCCGCTGTAAACAGAACTGAATTTGGACCCCGAG 2059
QY 37 uThrThrAspSerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuAr 57
Db 2060 CCAGTGGGCTCAGCCTGTATCCACAGCACTTTGGCAGGCCAGGTGGAGGATCACTTAG 2119
QY 57 gSerThrAspLeuLeuProTyrTyrLysIleAspAlaLeuCysLys 72
Db 2120 AGCTAGGAGTTTGAGACCAGCCTGG-----GCAATATGCAAG 2156

Search completed: January 15, 2006, 15:14:55
Job time : 177 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 06:11:50 ; Search time 3032 Seconds

(without alignments)
1188.194 Million cell updates/sec

Title: US-10-751-235-1

Perfect score: 393

Sequence: 1 LQPYAEDSANNWEAKPSQM.....STDLLPYMKIDALCKIVPRQ 77

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10751235/runat_12012006_120915_2527/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUPPIX=p2n.rst -MINMATCH=0 -L-LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blcosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235@cgn_1_1_5315 @runat_12012006_120915_2527 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	100.0	1664	CNS0A5YJ	BX822288 Arabidops
2	393	100.0	1709	CNS0A665	BX822785 Arabidops
3	356	90.6	858	CO072553	CO072553 GR_Ea31M
4	346	88.0	736	CW839414	CW839414 GT8284.Ds
5	343	87.3	531	BES52887	BES52887 946086H10
6	341	86.8	914	CK290930	CK290930 EST753644
7	337	85.8	840	BG643819	BG643819 EST512013

8	335	85.2	480	1	AJ486563
9	335	85.2	534	6	CA742365
10	335	85.2	596	1	AJ432371
11	335	85.2	600	5	BU987393
12	329.5	83.8	570	6	BH584135
13	327	83.2	657	6	CB255014
14	326	83.0	668	8	DR952300
15	326	83.0	886	8	DR923995
16	324	82.4	469	8	CX541939
17	324	82.4	652	2	BI263819
18	320.5	81.6	708	9	BH931006
19	319	81.2	626	5	BO971938
20	317	80.7	897	8	CV886111
21	307	78.1	693	10	CW166683
22	295	75.1	704	10	CZ710414
23	292	74.3	650	10	CZ799589
24	291	74.0	689	5	BQ862275
25	238	60.6	357	1	AJ615653
26	231	58.8	902	9	BZ822133
27	229	58.3	720	9	CC618386
28	229	58.3	891	10	CG056732
29	227	57.8	662	9	CC394464
30	225	57.3	932	10	CG056734
31	224	57.0	229	3	BI974897
32	224	57.0	1042	10	CW000851
33	218	55.5	860	10	CZ321241
34	214	54.5	679	7	CV538441
35	210	53.4	598	3	BM003139
36	207	52.7	474	3	BJ480615
37	207	52.7	598	3	BJ234910
38	207	52.7	634	8	CX023870
39	207	52.7	636	1	AV941342
40	207	52.7	697	1	AV939356
41	205	52.2	495	10	CG429086
42	205	52.2	571	3	BJ471574
43	205	52.2	700	6	CD053804
44	203	51.7	442	10	CG886172
45	201	51.1	905	10	CL426941

ALIGNMENTS

RESULT 1	CNS0A5YJ	1664 bp	mRNA	linear	HTC 06-FEB-2004
CNS0A5YJ	Arabidopsis thaliana Full-length cDNA Complete sequence from clone				
LOCUS	GLT2B2E04 of Flowers and buds of strain col-0 of Arabidopsis				
DEFINITION	thaliana (thale cress).				
ACCESSION	BX822288				
VERSION	BX822288.1	GI:42464384			
KEYWORDS	HTC; GSLT cDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 1664)				
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1664)				
AUTHORS	Genoscope.				
JOURNAL	Direct Submission				
	Submitted (18-NOV-2003) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.				

Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

source
1..1664
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTFB262B04"
/tissue_type="Flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..1664
/gene="At3g53130"

ORIGIN

Alignment Scores:
Pred. No.: 3,248-45 Length: 1664
Score: 393.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 4

US-10-751-235-1 (1-77) x CNS0A5YJ (1-1664)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
DB 624 TTGCAGCCTTATGCAGAGACGGAAGTCTGTGAATATGGAAGCGAGTCTCTCAGATG 683
QY 21 ThrLeuAppValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 684 ACACCTTGATGTCATTGGGTGTCTCTTTTAACTACAAATTCGATTCCTTGACTACTGAT 743
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 744 AGTCCTGTCATTGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGAT 803
QY 61 LeuLeuProTyrTriPylsIleAspAlaLeuCysLysIleValProArgGln 77
DB 804 CTTCTGCCATATGGNAGATCGATGATTTGTGAAGATAGTCCCGAGACAG 854

RESULT 2
CNS0A665 1709 bp mRNA linear HTC 06-FEB-2004
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTFB64ZA07 of Flowers and buds of strain col-0 of Arabidopsis
thaliana (Chale cress).

ACCESSION
BX822785

VERSION
BX822785.1 GI:42464524

KEYWORDS
HTC; GSLT cDNA.

SOURCE
Arabidopsis thaliana (thale cress)

ORGANISM
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1709)

Castelli, V., Aury J.M., Jaillon O., Wincker P., Clepet C.,

Menard M., Cruaud C., Quetier P., Scarpelli C., Schachter V.,

Temple, G., Caboche M., Weissenbach J., and Salanoubat M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished

2 (bases 1 to 1709)

Genoscope.

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
Genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

source

1..1709
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTFB64ZA07"
/tissue_type="Flowers and buds"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..1709
/gene="At3g53130"

gene

ORIGIN

Alignment Scores:
Pred. No.: 3,358-45 Length: 1709
Score: 393.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 4

US-10-751-235-1 (1-77) x CNS0A665 (1-1709)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
DB 613 TTGCAGCCTTATGCAGAGACGGAAGTCTGTGAATATGGAAGCGAGTCTCTCAGATG 672
QY 21 ThrLeuAppValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 673 ACACCTTGATGTCATTGGGTGTCTCTTTTAACTACAAATTCGATTCCTTGACTACTGAT 732
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 733 AGTCCTGTCATTGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGAT 792
QY 61 LeuLeuProTyrTriPylsIleAspAlaLeuCysLysIleValProArgGln 77
DB 793 CTTCTGCCATATGGNAGATCGATGATTTGTGAAGATAGTCCCGAGACAG 843

RESULT 3

CO072553

LOCUS

DEFINITION

GR_Ea31M14.r GR_Ea Gossypium raimondii cDNA clone GR_Ea31M14 3',
mRNA sequence.

ACCESSION
CO072553

VERSION
CO072553.1 GI:48742034

KEYWORDS
EST.

SOURCE
Gossypium raimondii

ORGANISM
Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

1 (bases 1 to 858)

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,

Udall, J. A., Rapp, R. A., Mendel, J. F., Rao, K., Soderlund, C. and

Wing, R. A.

Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 946086 row: H column: 10.

FEATURES

source
 1. 531
 /location/Qualifiers
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="OH43"
 /db_xref="taxon:4577"
 /tissue_type="tassel"
 /dev_stage="just after the transition from vegetative to inflorescence development"
 /lab_host="xLOUR"
 /clone_lib="946" - tassel primordium prepared by Schmidt lab
 /note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI; Site 2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA library in HybriZAP. Sample insert size range was 350 bp to 3 Kb with a 1 Kb average."

ORIGIN

Alignment Scores:
 Pred. No.: 1.12e-38 Length: 531
 Score: 343.00 Matches: 65
 Percent Similarity: 92.21% Conservative: 6
 Best Local Similarity: 84.42% Mismatches: 6
 Query Match: 87.28% Indels: 0
 Gaps: 0

US-10-751-235-1 (1-77) x BE552887 (1-531)

QY 1 LeuGlnProTyrrAlaGluAspGlySerAlaValaenMetGluAlaLysPheSerGlnMet 20
 DB 1 CTTGAGCATATGCTTTCAGTGGGACCTGTCATATGGAAGCGAGGTTTCTCAGTTG 60
 QY 21 ThrLeuAppValIleGlyLeuSerLeuPheAenTyrrAenPheAenSerLeuThrAsp 40
 DB 61 ACATTTGGATGTCATGTTTATCATTTGTTCAACTACAAATTTTGAATCCCTCAACACAGAT 120
 QY 41 SerProValIleGluAlaValTyrrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 DB 121 AGTCTGTGATGATGCTGTTTATCTGCACTCAAGAGACGAGCTGCTTCTACAGAT 180
 QY 61 LeuLeuProTyrrTriPlysIleAspAlaLeuCysLysIleValProArgGln 77
 DB 181 CTTTGGCATATGGAAGGTTGGTTCTTGTGCAAGATAATCCCAAGACAG 231

RESULT 6
 CK290930 914 bp mRNA linear EST 02-AUG-2004
 LOCUS EST753644 Nicotiana benthamiana mixed tissue cDNA library,
 DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NEMBY36 5' end, mRNA sequence.

ACCESSION CK290930
 VERSION CK290930.1 GI:39870877
 KEYWORDS EST.
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Nicotiana.

1 (bases 1 to 914)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,

Staskawicz, B., Jin, H., and Baker, B.

Generation of EST sequences from Nicotiana benthamiana

Unpublished (2003)

Other ESTs: EST753645

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics
 Institute via <http://genome.arizona.edu/orders/>.
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source
 1. 914
 /location/Qualifiers
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NEMBY36"
 /tissue_type="abiotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Nicotiana benthamiana
 tissues that include callus, roots from liquid culture
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;
 Xanthomonas campestris pv campestris 12 hr, 18hr;
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
 campestris pv vesicatoria 18hr). RNA was isolated from
 these tissues and pooled in approximately equal molar
 amounts."

ORIGIN

Alignment Scores:
 Pred. No.: 4.43e-38 Length: 914
 Score: 341.00 Matches: 66
 Percent Similarity: 93.51% Conservative: 6
 Best Local Similarity: 85.71% Mismatches: 5
 Query Match: 86.77% Indels: 0
 Gaps: 0

US-10-751-235-1 (1-77) x CK290930 (1-914)

QY 1 LeuGlnProTyrrAlaGluAspGlySerAlaValaenMetGluAlaLysPheSerGlnMet 20
 DB 481 CTTACACCTGATGCGAATTTCTGGCTCTGCACTTAATATGAGGCAAGGTTTCTCACTA 540
 QY 21 ThrLeuAppValIleGlyLeuSerLeuPheAenTyrrAenPheAenSerLeuThrAsp 40
 DB 541 ACACCTTGATGATGTTTGGCTTTCACCTTCAATTAACAATTTTGAATCCCTTACTACTGAC 600
 QY 41 SerProValIleGluAlaValTyrrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 DB 601 AGTCACGTTATTGAAGCAGTTTACACTGCTGTTAAAGAGAGCAACTCCGTTCAACTGAT 660
 QY 61 LeuLeuProTyrrTriPlysIleAspAlaLeuCysLysIleValProArgGln 77
 DB 661 CTGTTGCCGCTATTGGCAGATCAAGCCTTATGTAAGGTCAATCCACGACAA 711

RESULT 7

BG643819

LOCUS BG643819

DEFINITION ESTS12013 tomato shoot/meristem Lycopersicon esculentum cDNA clone

CTOP32N21 5' sequence, mRNA sequence.

ACCESSION BG643819

VERSION BG643819.1

KEYWORDS GI:13779044

SOURCE EST.

ORGANISM Lycopersicon esculentum (Solanum lycopersicum)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 840)

van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,

Hansen, C., Roming, C. and Tankalev, S.

Generation of ESTs from tomato shoot/meristem tissue

Unpublished (2001)

Contact: CUGI

COMMENT Clemson University Genomics Institute

Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES

source

1. .840
/organism="Lycopersicon esculentum"
/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="CTOF32N21"

/tissue type="shoot/meristem"

/dev stage="developing shoots from 4-6wks old plants"

/lab_host="SOLr"

/clone_lib="tomato shoot/meristem"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; Small expanding leaves from the growing tip were

taken from greenhouse plants (4-6wks old TA496). Tissue

was immediately frozen in liquid nitrogen."

ORIGIN

Alignment Scores:

Pred. No.: 1.49e-37 Length: 840
Score: 337.00 Matches: 66
Percent Similarity: 92.21% Conservative: 5
Best Local Similarity: 85.71% Mismatches: 6
Query Match: 85.75% Indels: 0
DB: 2 Gaps: 0

US-10-751-235-1 (1-77) x BG643819 (1-840)

QY 1 LeuGinProTyAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
DB 32 CTTTACCTGATGCAATTTCTGGCTCTGCAGTGAATATGAGCGCAAGTTTCTCAACTA 91
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyAsnPheAspSerLeuThrThrAsp 40
DB 92 ACACCTGATGTTATGGCTTGCACCTCTCAATTACATTTGATTCCTTACTACTGAC 151
QY 41 SerProValIleGluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 152 AGTCAGTTATTGTCAGTTTACACTGCCTAAAAGAGCAGAACTCGCTTCAACTGAT 211
QY 61 LeuLeuProTyTrrLysIleAspAlaLeuCyLysIleValProArgGln 77
DB 212 TTGTTGCCATATGGCAGATCAAAAGCTTTATGTAAGTTTCATCCCGACAA 262

RESULT 8

AJ486563

LOCUS AJ486563 S00011 Hordeum vulgare cDNA clone S0001100041C11F1, mRNA
DEFINITION sequence.

AJ486563

ACCESSION AJ486563.1 GI:21202517

VERSION

KEYWORDS

SOURCE

ORGANISM

Hordeum vulgare

Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Hordeum.

1 (bases 1 to 480)

Saren, A.-M., Tanskanen, J., Paulin, L. and Schulman, A.H.

Barley EST's

Unpublished (2002)

Contact: Schulman AH

Institute of Biotechnology

University of Helsinki

P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014,

Finland.

FEATURES

source

1. .480

/organism="Hordeum vulgare"

/mol_type="mRNA"

/db_xref="taxon:4513"

/clone="S0001100041C11F1"
/dev stage="Developing seed"
/clone_lib="S00011"
/note="12,15,18 days after pollination"

ORIGIN

Alignment Scores:

Pred. No.: 1.38e-37 Length: 480
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 1 Gaps: 0

US-10-751-235-1 (1-77) x AJ486563 (1-480)

QY 1 LeuGinProTyAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
DB 216 CTCGAGACATATGCTTTGAGCGGTGAACCTGTTATATGGAAGCGAGATTTCTCAATG 275
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyAsnPheAspSerLeuThrThrAsp 40
DB 276 ACACCTAGATGCTGTTGTTGTTGTTTCAACTCAACTTTGATTCCTCCTCAGATCAGAT 335
QY 41 SerProValIleGluAlaValTyThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 336 AGTCCTGTTATTGATGCTGTTTACACCGCACCTGAAAGAGCAGAGGCTGCTTCTCAGAT 395
QY 61 LeuLeuProTyTrrLysIleAspAlaLeuCyLysIleValProArgGln 77
DB 396 CTTTACCATCTGCGAGATTTGTTGCTGCGAAGATTTCTTAGACAG 446

RESULT 9

CA742365

LOCUS

DEFINITION

CA742365

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Triticum aestivum

Triticum aestivum (bread wheat)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Triticum.

1 (bases 1 to 534)

Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,

Miso, G., Caraher, N. and Hanafey, M.K.

DuPont Wheat cDNA Sequence

Unpublished (2002)

Contact: Scott V. Tingey

Crop Genetics

E. I. DuPont de Nemours and Company

1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA

Tel: 302-631-2607

Fax: 302-631-2607

Email: Scott.V.Tingey@USA.dupont.com

Seq primer: M13.

Location/Qualifiers

1.534

/organism="Triticum aestivum"

/mol_type="mRNA"

/db_xref="taxon:4565"

/clone="wfl1c.pk001.i13"

/tissue type="leaf"

/lab_host="DH10B"

/clone_lib="wfl1c"

/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:

XhoI; Wheat (Triticum aestivum, Hi Line) flag leaf"

ORIGIN

Alignment Scores:

Pred. No.: 1.59e-37 Length: 534

```

Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 6 Gaps: 0

US-10-751-235-1 (1-77) x CA742365 (1-534)

Qy 1 LeuGlnProTyTrpAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 82 CTCGAGACTTATGCTTTCGAGCGTGAACCTGTTAATATGGAAGCGAGGTTTCTCAAAATG 141
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrrAsnPheAspSerLeuThrThrAsp 40
Db 142 ACATTAGATGTGATGGTGTATTCCTTGTTCACACTCAACTTTCGATTCCTCCTCAGAT 201
Qy 41 SerProValIleGluAlaValTyrrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 202 AGTCCTGTTATGATGCTGTTCACACTGCATCAAGAAGCTGAGGCTCGTTCACAGAT 261
Qy 61 LeuLeuProTyTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 262 CTTTACCATACTGCAGATGATTTGCTGTGCAAGATTTGTTCTCCTAGACAG 312

RESULT 10
AJ432371 596 bp mRNA linear EST 15-MAR-2002
LOCUS AJ432371 S00011 Hordeum vulgare cDNA clone S0001100019A01F1, mRNA
DEFINITION
ACCESSION AJ432371
VERSION AJ432371.1 GI:19520823
KEYWORDS EST.
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 596)
Saren, A.-M., Tanskanen, J., Paulin, L. and Schulman, A.H.
Barley EST's
Unpublished (2002)
Contact: Schulman AH
Institute of Biotechnology
University of Helsinki
P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014,
Finland.

FEATURES
source
1..596
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="taxon:4513"
/clone="S0001100019A01F1"
/dev_stage="Developing seed"
/clone_lib="S00011"
/notes="12,15,18 days after pollination"

ORIGIN
Alignment Scores:
Pred. No.: 1 84e-37 Length: 596
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 1 Gaps: 0

US-10-751-235-1 (1-77) x AJ432371 (1-596)

Qy 1 LeuGlnProTyTrpAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 226 CTCGAGACATATGCTTTCGAGCGTGAACCTGTTAATATGGAAGCGAGATTTTCTCAAAATG 285
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrrAsnPheAspSerLeuThrThrAsp 40

```

```

Db 286 AACTAGATGTGATTTGTTGTTCTTGTTCACACTACACTTGTGATTCCTCCTCAGAT 345
Qy 41 SerProValIleGluAlaValTyrrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 346 AGTCCTGTTATGATGCTGTTTACACCGCATCTGAAGAAGCAGAGGCTCGTTCACAGAT 405
Qy 61 LeuLeuProTyTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 406 CTTTACCATACTGCAGATGATTTGCTGTGCAAGATTTGTTCTCCTAGACAG 456

RESULT 11
BU987393 600 bp mRNA linear EST 22-OCT-2002
LOCUS BU987393 HF14K04r Hordeum vulgare subsp. vulgare cDNA clone HF14K04
DEFINITION
ACCESSION BU987393
VERSION BU987393.1 GI:24238339
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 600)
Radchuk, V., Zhang, H., Weschke, W., Potokina, E. and Wobus, U.
Barley ESTs from developing seeds
Unpublished (2002)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 600 Std Error: 0.00
Plate: 14 row: K column: 4
Seq primer: M13rev.

FEATURES
source
1..600
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Barke"
/sub_species="vulgare"
/db_xref="GABI:247336"
/db_xref="taxon:112509"
/clone="HF14K04"
/tissue_type="caryopsis"
/dev_stage="developing caryopsis, 16-25 DAF (days after
flowering)"
/lab_host="XL10-Gold"
/clone_lib="HF"
/notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
cDNA); Site 2: XhoI (3'-end of cDNA); developing
caryopsis, 16-25 DAF(days after flowering) Due to a cloning
artefact caused by the kit, in most cases the EcoRI site
is NOT present, as well as the EcoRIadapter used for
cloning. To excise the insert, restriction sites upstream
EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
due to the cloning system used Blue/White selection for
recombinants is not 100% reliable.Average insert size is
940 bp"

ORIGIN
Alignment Scores:
Pred. No.: 1 85e-37 Length: 600
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservative: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 5 Gaps: 0

US-10-751-235-1 (1-77) x BU987393 (1-600)

```

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
 Db 176 CTCGACATATGCTTTGAGCGGTGAACCTGTTAATATGGAAGCGAGATTTTCTCAATG 235
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 Db 236 ACACATAGATGATGTTGTTGCTTTGTTCACTACACTTTGATTCCTCAGATCAGAT 295
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 Db 296 AGTCCTGTTATTGATGCTGTTTACACCGCACTGMAAGACGAGAGGCTGTTCTACAGAT 355
 QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
 Db 356 CTTTACCATACTGGCAGATGATTGCTGTGCAAGATTGTTCTCAGACAG 406

RESULT 12
 BH584135/c
 LOCUS BH584135
 DEFINITION BOHAD90TF BOHA Brassica oleracea genomic clone BOHAD90, genomic survey sequence.
 ACCESSION BH584135
 VERSION BH584135.1 GI:17836592
 KEYWORDS GSS
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 570)
 AUTHORS Ayele, M., Haas, B.J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S., Utterback, T.R., Wortman, J.R., White, O.R. and Town, C.D.
 TITLE Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
 JOURNAL Genome Res. 15 (4), 487-495 (2005)
 PUBMED 15805490
 COMMENT Other_GSSs: BOHAD90TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TP
 Class: sheared ends.

FEATURES
 source
 1..570
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone="BOHAD90"
 /clone_lib="BOHA"
 /note="Vector: pHOS1; Site 1: BatXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Alignment Scores:
 Pred. No.: 1.07e-36 Length: 570
 Score: 329.50 Matches: 71
 Percent Similarity: 71.84% Conservativeness: 3
 Best Local Similarity: 68.93% Mismatches: 3
 Query Match: 83.84% Indels: 26
 DB: 9 Gaps: 1

US-10-751-235-1 (1-77) x BH584135 (1-570)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
 Db 442 CTCGACCTTATGATGAGTGGAGGAAAGCTGTGAACATGGAGAGAGATTTCTCAGTTA 383
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40

Db 382 ACTCTCGAGCGATGATGGCTTATCTCTTTTAACTACAACCTTTGATTCCTCACTACCGAT 323
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 Db 322 AGTCCTGTCATCGAAGCTGTTTACACTGCTCTCAAGAAGCTGAGCTTGCTTCTACTGAT 263
 QY 61 LeuLeuProTyrTrpLys-----
 Db 262 ATTACCCCTATTGGAAGCGAAGTTTGTGTTCTTCTGTGGCGCAGACTTGGATTCT 203
 QY 67 -----leaspAlaLeuCysLysIleValPr 75
 Db 202 TTTTGAGCCTTGGTGTGTTTCTTGTTCAGATCGATCGTGTGTGAAGATAGTTC 143
 QY 75 oArgGln 77
 Db 142 GAGACAA 136

RESULT 13
 CB255014
 LOCUS

DEFINITION 54-E018363-019-007-L14-T7R MP1Z-ADIS-019 Arabidopsis thaliana cDNA clone MP1Zp768L147Q 5-PRIME, mRNA sequence.
 ACCESSION CB255014
 VERSION CB255014.1 GI:56915639
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 657)
 AUTHORS Jakob, M., Stracke, R., Soerensen, T.R. and Weisshaar, B.
 TITLE Arabidopsis thaliana cDNA library enriched in transcription factors
 JOURNAL Unpublished (2003)
 COMMENT Contact: Weisshaar B

ADIS DNA core facility at MP1Z
 Max-Planck-Institute for Plant Breeding Research
 Carl-von-Linne Weg 10, 50829 Koeln, Germany
 Fax: 00492215062851
 Email: weisshaar@mpiz-koeln.mpg.de
 Insert Length: 657 Std Error: 0.00
 Plate: 7 row: L column: 14
 Seq primer: T7R; CTAATACGACTCACTATAGGGA.

FEATURES
 Location/Qualifiers
 1..657
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /cultiivar="At7"
 /db_xref="GABI:597448"
 /db_xref="taxon:3702"
 /clone="MP1Zp768L147Q"
 /tissue_type="hypocotyl"
 /dev_stage="tissue culture"
 /lab_host="E. coli DH5alpha mcr"
 /clone_lib="MP1Z-ADIS-019"
 /note="Vector: pSPORT1; Site 1: NotI primer adapter; Site 2: SalI primer adapter; RNA from cellculture (At7) 5 days after inoculation treated with 0.002 mM cycloheximide for 2 h in the dark. Sequence submission managed by RZPD/GABI-Primary database: http://gabi.rzpd.de. This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."

ORIGIN

Alignment Scores:
 Pred. No.: 2.96e-36 Length: 657
 Score: 327.00 Matches: 68
 Percent Similarity: 97.18% Conservativeness: 1
 Best Local Similarity: 95.77% Mismatches: 0
 Query Match: 83.21% Indels: 2
 DB: 6 Gaps: 0

US-10-751-235-1 (1-77) x CB255014 (1-657)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValaAenMetGluAlaLysPheSerGlnMet 20
 Db 447 TTGCAGCCTTATGCAGNAGACGGAAGTCTGTGAATATGGAAGCGAAGTCTCTCAGATG 506
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAenTyrAenPheAenSerLeuThrThrAsp 40
 Db 507 ACACCTGATGTCATTTGGTGTCTCTTTTAACTACAAATTTTCGATTTCTTGACTACTGAT 566
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAla--GluLeuArgSerThrA 60
 Db 567 AGTCTGTGTCATGGAAGCTGTGTACACTGCTCTTAAAGAAGCTTGAGCTTCTGTTCTACTG 626
 QY 60 spLeuLeuProTyrTrpLysIleAspAla 69
 Db 627 ATCTTCTGCCATATTGGAAGATCGATGCA 655
 DR952300 668 bp mRNA linear EST 02-AUG-2005
 EST1143839 Aquilegia cDNA library Aquilegia formosa x Aquilegia
 pubescens cDNA clone COLS013, mRNA sequence.
 DR952300
 DR952300.1 GI:71721663
 EST.
 Aquilegia formosa x Aquilegia pubescens
 Aquilegia formosa x Aquilegia pubescens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 Ranunculaceae; Aquilegia.
 1 (bases 1 to 668)
 Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,
 Nordborg, M. and Tomkins, J.
 Generation of ESTs from Aquilegia
 Unpublished (2005)
 Other ESTs: EST1143838
 Contact: Scott Hodges
 Department of Ecology, Evolution and Marine Biology
 University of California, Santa Barbara
 Santa Barbara, CA 93106, USA
 Tel: 805 893 7813
 Fax: 805 893 4724
 Email: hodges@lifesci.ucsb.edu
 Seq primer: M13 Reverse.

FEATURES
 source

1..668
 /organism="Aquilegia formosa x Aquilegia pubescens"
 /mol_type="mRNA"
 /db_xref="taxon:338618"
 /clone="COLS013"
 /tissue_type="mixed shoot and floral apical meristems,
 flower buds, leaves and roots"
 /lab_host="DH10B T1 (T1 and T5 phage resistance)"
 /clone_lib="Aquilegia cDNA library"
 /note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
 F2, F3, and F4 lines of Aquilegia formosa x A. pubescens
 were grown from seed in greenhouses at UC Santa Barbara.
 From these plants three sets of tissue were collected: 1)
 Small flower buds (<10 mm) and very young inflorescences
 (71 & 29% by weight respectively), 2) Medium (7-20 mm) and
 large (at or near anthesis) flower buds (65 & 35% by
 weight respectively) and 3) Shoot apical meristems. A
 fourth set of tissue was collected from plants of A.
 formosa. These plants were grown from seed in sand and at
 approximately 1 month root tissue and leaf tissue of
 various developmental stages were collected (84 & 16% by
 weight respectively). Total RNA was extracted from each
 set of tissue and pooled in the following proportions:
 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled
 total RNA, mRNA was extracted and enriched for full-length
 messages and then normalized with proprietary methods by

Invitrogen."

ORIGIN

Alignment Scores: Length: 668
 Pred. No.: 4,21e-36 Matches: 63
 Score: 326.00 Conservat: 6
 Percent Similarity: 89.61% Mismatches: 8
 Best Local Similarity: 81.82% Indels: 0
 Query Match: 82.95% Gaps: 0
 Db: 8

US-10-751-235-1 (1-77) x DR952300 (1-668)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValaAenMetGluAlaLysPheSerGlnMet 20
 Db 428 CTTAGACCTGATGCTGTGAGTGGTCTCTCTGAATATGGAAGCAAGTTTCGAGTTA 487
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAenTyrAenPheAenSerLeuThrThrAsp 40
 Db 488 ACTTAGATGTTATAGGTTTGTCACTGTTCAATATATATTTTGAATTCATTAATCGGAT 547
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 Db 548 AGTCTGTGTCATGGAAGCTGTGTACACTGCTCTTAAAGAAGCTTGAGCTTCTGTTCTACTG 626
 QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
 Db 608 CTTTACCATATTGGAAGATTAACTTTGTGCAAGATAAATCCGAGGCAA 658

RESULT 15

DR923995 886 bp mRNA linear EST 02-AUG-2005
 EST1115534 Aquilegia cDNA library Aquilegia formosa x Aquilegia
 pubescens cDNA clone COLM240, mRNA sequence.
 DR923995
 DR923995.1 GI:71693358
 EST.
 Aquilegia formosa x Aquilegia pubescens
 Aquilegia formosa x Aquilegia pubescens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 Ranunculaceae; Aquilegia.
 1 (bases 1 to 886)
 Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E.,
 Nordborg, M. and Tomkins, J.
 Generation of ESTs from Aquilegia
 Unpublished (2005)
 Other ESTs: EST1115533
 Contact: Scott Hodges
 Department of Ecology, Evolution and Marine Biology
 University of California, Santa Barbara
 Santa Barbara, CA 93106, USA
 Tel: 805 893 7813
 Fax: 805 893 4724
 Email: hodges@lifesci.ucsb.edu
 Seq primer: M13 Reverse.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..886
 /organism="Aquilegia formosa x Aquilegia pubescens"
 /mol_type="mRNA"
 /db_xref="taxon:338618"
 /clone="COLM240"
 /tissue_type="mixed shoot and floral apical meristems,
 flower buds, leaves and roots"
 /lab_host="DH10B T1 (T1 and T5 phage resistance)"
 /clone_lib="Aquilegia cDNA library"
 /note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
 F2, F3, and F4 lines of Aquilegia formosa x A. pubescens
 were grown from seed in greenhouses at UC Santa Barbara.
 From these plants three sets of tissue were collected: 1)
 Small flower buds (<10 mm) and very young inflorescences
 (71 & 29% by weight respectively), 2) Medium (7-20 mm) and
 large (at or near anthesis) flower buds (65 & 35% by

weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

ORIGIN

Alignment Scores:

Pred. No.:	6.12e-36	Length:	886
Score:	326.00	Matches:	63
Percent Similarity:	89.61%	Conservative:	6
Best Local Similarity:	81.82%	Mismatches:	8
Query Match:	82.95%	Indels:	0
DB:	8	Gaps:	0

US-10-751-235-1 (1-77) x DR923995 (1-886)

Qy	1	LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet	20
Db	646	CTTAGACCTGATGCTGAGTGGTCTCTGTGAATATGGAGGCAAGTTTCGCAGTTA	705
Qy	21	ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp	40
Db	706	ACTTTAGATGTTATAGGTTTGTCACGTTCATTAATTTTGATTCACTTAATCGCGAT	765
Qy	41	SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp	60
Db	766	AGTCCTGTTATTGATGCAGTATATCTGCATTAAAGAGGACAGGCTCGCTCAACAGAT	825
Qy	61	LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln	77
Db	826	CTTCTACCATATTGGAAGATTAACTTTTGTGCAAGATAATTCCGAGGCAA	876

Search completed: January 15, 2006, 15:11:57
Job time : 3038 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 15, 2006, 13:12:22 ; Search time 615 Seconds
(without alignment)
1035.353 Million cell updates/sec

Title: US-10-751-235-1
Perfect score: 393
Sequence: 1 LQYABDGSVAVMEAKFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2_1/USPTO.spool/US10751235/runat 12012006 120916 2564/app query.fasta_1.263
-DB=Published Applications NA Main -QFWT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCK=100 -LONGLOG
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235.ecgn 1.1 1026 @runat 12012006 120916 2564 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUB -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:
1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgm2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgm2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgm2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgm2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	100.0	3377	8	US-10-739-930-773
2	346	88.0	2467	9	US-10-751-235-5
3	346	88.0	2467	9	US-10-751-235-7
4	346	88.0	4170	9	US-10-751-235-6
5	343	87.3	531	9	US-10-751-235-26
6	343	87.3	1804	7	US-10-425-114-32608
7	343	87.3	2537	8	US-10-425-115-148190

8	337	85.8	839	9	US-10-751-235-25	Sequence 25, Appl
9	335	85.2	1086	7	US-10-751-235-24	Sequence 24, Appl
10	335	85.2	1118	7	US-10-437-963-15200	Sequence 15200, A
11	335	85.2	1638	9	US-10-751-235-23	Sequence 23, Appl
12	335	85.2	1686	9	US-10-751-235-22	Sequence 22, Appl
13	319	81.2	624	9	US-10-751-235-27	Sequence 27, Appl
14	224	57.0	229	7	US-10-424-599-9451	Sequence 9451, Ap
15	210	53.4	1101	9	US-10-751-235-48	Sequence 48, Appl
16	207	52.7	1031	9	US-10-751-235-46	Sequence 46, Appl
17	207	52.7	1527	9	US-10-751-235-44	Sequence 44, Appl
18	206	52.4	1767	7	US-10-437-963-37324	Sequence 37324, A
19	206	52.4	1899	9	US-10-751-235-43	Sequence 43, Appl
20	202	51.4	2239	7	US-10-425-114-78	Sequence 78, Appl
21	202	51.4	2404	8	US-10-425-115-59515	Sequence 59515, A
22	200	50.9	1731	9	US-10-751-235-55	Sequence 55, Appl
23	200	50.9	1909	7	US-10-424-599-110390	Sequence 110390, A
24	199	50.6	2181	7	US-10-437-963-53676	Sequence 53676, A
25	199	50.6	2211	8	US-10-425-115-33709	Sequence 33709, A
26	198	50.4	1788	9	US-10-751-235-41	Sequence 41, Appl
27	198	50.4	2057	9	US-10-751-235-40	Sequence 40, Appl
28	196	49.9	404	6	US-10-174-693-156	Sequence 156, App
29	196	49.9	1278	9	US-10-751-235-45	Sequence 45, Appl
30	196	49.9	1734	9	US-10-751-235-53	Sequence 53, Appl
31	193	49.1	1596	7	US-10-437-963-15203	Sequence 15203, A
32	193	49.1	1737	9	US-10-751-235-47	Sequence 47, Appl
33	191.5	48.7	1926	9	US-10-751-235-54	Sequence 54, Appl
34	179	45.5	5071	9	US-10-751-235-42	Sequence 42, Appl
35	177.5	45.2	1980	9	US-10-751-235-57	Sequence 57, Appl
36	96	24.4	3219	6	US-10-156-761-567	Sequence 567, App
37	96	24.4	9025608	6	US-10-156-761-1	Sequence 1, Appl1
38	88	22.4	1401	9	US-10-869-813-19	Sequence 19, Appl
39	88	22.4	3162	9	US-10-869-813-107	Sequence 107, App
40	87	22.1	1216	8	US-10-425-115-59514	Sequence 59514, A
41	82	20.9	1401	9	US-10-869-813-3	Sequence 3, Appl1
42	82	20.9	3183	9	US-10-869-813-105	Sequence 105, App
43	80.5	20.5	726	3	US-09-974-300-2232	Sequence 2232, Ap
44	76	19.3	1392	9	US-10-869-813-1	Sequence 1, Appli
45	76	19.3	3147	9	US-10-869-813-103	Sequence 103, App

ALIGNMENTS

RESULT 1

US-10-739-930-773
; Sequence 773, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 773
; LENGTH: 3377
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER3001_1
US-10-739-930-773

Alignment Scores:
Pred. No.: 8.9e-51 Length: 3377
Score: 393.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-751-235-1 (1-77) x US-10-739-930-773 (1-3377)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20

Db 676 TTGACGCTTATGCAGACGAGAGTCTGTGATATATGGAGCGAAGTTCTCTCAGATG 735
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 736 ACACCTTGATGTCATTTGGTGTCTCTTTTAACTACAAATTTTCGATTTCTTGACTACTGAT 795
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 796 AGTCCTGTCAATGAAGCTGTTTACACTGCTCTTAAAGAGAGCTGAGCTTCGTTCTACTGAT 855
Qy 61 LeuLeuProTyrTrpLysAlaLeuCysValIleValProArgGln 77
Db 856 CTTCTGCCATATTGGAAGATCGATGTCATTTGTGTAAGATAGTCCCGAGACAG 906

RESULT 2

US-10-751-235-5
; Sequence 5, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-751-235-5

Alignment Scores:
Pred. No.: 1.65e-43 Length: 2467
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservative: 0
Best Local Similarity: 71.03% Mismatches: 1
Query Match: 88.04% Indels: 31
Gaps: 1

US-10-751-235-1 (1-77) x US-10-751-235-5 (1-2467)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 747 TTGCAGCCTTATGCAGACGAGAGTCTGTGATATATGGAGCGAAGTTCTCTCAGATG 806
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 807 ACACCTTGATGTCATTTGGTGTCTCTTTTAACTACAAATTTTCGATTTCTTGACTACTGAT 866
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 867 AGTCCTGTCAATGAAGCTGTTTACACTGCTCTTAAAGAGAGCTGAGCTTCGTTCTACTGAT 926
Qy 61 LeuLeuProTyrTrpLysAlaLeuCysValIleValProArgGln 77
Db 927 CTTCTGCCATATTGGAAGAGC-AGTTTCTGTGTTTTTCTGTGGTTTGTGATTTGTTG 985
Qy 68 -----AspAlaLeu 70
Db 986 GAACAATTGATTTCTTTAAATTGAGAGGTTTGGTTGTTTTTTTTCAGATCGATGATG 1045
Qy 71 CysLysIleValProArgGln 77
Db 1046 TGTAAAGATAGTCCCGAGACAG 1066

RESULT 3

US-10-751-235-7
; Sequence 7, Application US/10751235

; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-751-235-7

Alignment Scores:
Pred. No.: 1.65e-43 Length: 2467
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservative: 0
Best Local Similarity: 71.03% Mismatches: 1
Query Match: 88.04% Indels: 31
Gaps: 1

US-10-751-235-1 (1-77) x US-10-751-235-7 (1-2467)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 747 TTGCAGCCTTATGCAGACGAGAGTCTGTGATATATGGAGCGAAGTTCTCTCAGATG 806
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 807 ACACCTTGATGTCATTTGGTGTCTCTTTTAACTACAAATTTTCGATTTCTTGACTACTGAT 866
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 867 AGTCCTGTCAATGAAGCTGTTTACACTGCTCTTAAAGAGAGCTGAGCTTCGTTCTACTGAT 926
Qy 61 LeuLeuProTyrTrpLysAlaLeuCysValIleValProArgGln 77
Db 927 CTTCTGCCATATTGGAAGGC-AGTTTCTGTGTTTTTCTGTGGTTTGTGATTTGTTG 985
Qy 68 -----AspAlaLeu 70
Db 986 GAACAATTGATTTCTTTAAATTGAGAGGTTTGGTTGTTTTTTTTCAGATCGATGATG 1045
Qy 71 CysLysIleValProArgGln 77
Db 1046 TGTAAAGATAGTCCCGAGACAG 1066

RESULT 4

US-10-751-235-6
; Sequence 6, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 4170
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-751-235-6

Alignment Scores:
Pred. No.: 3.6e-43 Length: 4170
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservative: 0
Best Local Similarity: 71.03% Mismatches: 1
Query Match: 88.04% Indels: 31
DB: 9 Gaps: 1

US-10-751-235-1 (1-77) x US-10-751-235-6 (1-4170)

Qy 1 LeuGlnProTyrAlaGluaspGlySerAlaValaenMetGluAlaLysPheSerGlnMet 20
Db 1756 TTGCAGCCTTATGCAGAGAGGAGTGTGTGAATATGGAAGCGAGTCTCTCAGATG 1815

Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAenPheAspSerLeuThrAsp 40
Db 1816 ACATTGATGTCATGGGTGTCTCTTTTAACTACAATTTTCGATCTCTTGACTACTGAT 1875

Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 1876 AGTCTGTCATGAAGCTGTGTACACTGCTCTTAAAGAAGCTGAGCTTCTGTTCTACTGAT 1935

Qy 61 LeuLeuProTyrTrpLysIle----- 67
Db 1936 CTCTGCCATATGGAAGGC-AGGTTTCTGTGTGTTTTTCTGTGTTTGTGATGTGTG 1994

Qy 68 -----AspAlaLeu 70
Db 1995 GAACAATGGATCTCTGTTAATTGAGAGGTTTGTGTTGTTTTCAGATCGATGCAATTG 2054

Qy 71 CysLeuValIleValProArgGln 77
Db 2055 TGTAAGATAGTCCCGAGACAG 2075

RESULT 5
US-10-751-235-26
; Sequence 26, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Zea mays
US-10-751-235-26

Alignment Scores:
Pred. No.: 5.05e-44 Length: 531
Score: 343.00 Matches: 65
Percent Similarity: 92.21% Conservative: 6
Best Local Similarity: 84.42% Mismatches: 0
Query Match: 87.28% Indels: 0
DB: 9 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-26 (1-531)

Qy 1 LeuGlnProTyrAlaGluaspGlySerAlaValaenMetGluAlaLysPheSerGlnMet 20
Db 1 CTTGAGCCATATGCTTGGAGTGGGAACCTGTCAATATGGAAGCGAGGTTTCTCAGTTG 60

Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAenPheAspSerLeuThrAsp 40
Db 61 ACATTGATGTCATGGGTGTCTCTTTTAACTACAATTTTGTTCACATATTTTCCCTCACAACAGAT 120

Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 614 AGTCTGTCATGATGCTGTTTATCTGCACTCAAGAGAGCAGGCTTCTGTTCTACAGAT 673

Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 674 CTTTGGCATACTGGAAGTGGTTTCTTGTGCAAGATAATCCCAAGACAG 724

RESULT 7
US-10-425-115-148190
; Sequence 148190, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28

Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 121 AGTCTGTCATGATGCTGTTTATCTGCACTCAAGAGAGCAGGCTTCTGTTCTACAGAT 180

Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 181 CTTTGGCATACTGGAAGTGGTTTCTTGTGCAAGATAATCCCAAGACAG 231

RESULT 6
US-10-425-114-32608
; Sequence 32608, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32608
; LENGTH: 1804
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17009E12_PLI
US-10-425-114-32608

Alignment Scores:
Pred. No.: 3.1e-43 Length: 1804
Score: 343.00 Matches: 65
Percent Similarity: 92.21% Conservative: 6
Best Local Similarity: 84.42% Mismatches: 0
Query Match: 87.28% Indels: 0
DB: 7 Gaps: 0

US-10-751-235-1 (1-77) x US-10-425-114-32608 (1-1804)

Qy 1 LeuGlnProTyrAlaGluaspGlySerAlaValaenMetGluAlaLysPheSerGlnMet 20
Db 494 CTTGAGCCATATGCTTGGAGTGGGAACCTGTCAATATGGAAGCGAGGTTTCTCAGTTG 553

Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAenPheAspSerLeuThrAsp 40
Db 554 ACATTGATGTCATGGGTGTCTCTTAACTACAATTTTGTTCACATATTTTCCCTCACAACAGAT 613

Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 614 AGTCTGTCATGATGCTGTTTATCTGCACTCAAGAGAGCAGGCTTCTGTTCTACAGAT 673

Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 674 CTTTGGCATACTGGAAGTGGTTTCTTGTGCAAGATAATCCCAAGACAG 724

; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 148190
; LENGTH: 2537
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MWT4577_66678C.1
US-10-425-115-148190

Alignment Scores:
Pred. No.: 5, 15e-43 Length: 2537
Score: 343.00 Matches: 65
Percent Similarity: 92.21% Conservatives: 6
Best Local Similarity: 84.42% Mismatches: 6
Query Match: 87.28% Indels: 0
DB: 8 Gaps: 0

US-10-751-235-1 (1-77) x US-10-425-115-148190 (1-2537)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 747 CTTGAGCCATATGCTTTCAGTGGGGAACCTGTCAATATGGAAGCAGGTTTCTCAGTTG 806
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
Db 807 ACATTGGATGTGATTGGTTTATCATTTGTCAACTACAATTTTGATTCCTCCACACAGAT 866
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 867 AGTCCTGTGATGATGCTGTGTTTATCTGCACTCAAGAAGCAGAGCTTCGTTCTACAGAT 926
Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 927 CTTTGGCCATACATGGAAGTTGGTTCTGTGCAAGATAATCCCAAGACAG 977

RESULT 8

US-10-751-235-25
; Sequence 25, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 839
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-10-751-235-25

Alignment Scores:
Pred. No.: 8, 95e-43 Length: 839
Score: 337.00 Matches: 66
Percent Similarity: 92.21% Conservatives: 5
Best Local Similarity: 85.71% Mismatches: 6
Query Match: 85.75% Indels: 0
DB: 9 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-25 (1-839)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 31 CTTTACCTGATGCAATTTCTGGCTCTGCAGTGAATATGAGGCAAAAGTTTCTCAACTA 90
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
Db 91 ACATTGGATGTATGGCCCTTGCACTCTTCATTAACAATTTTGATTCCTACTACTGAC 150

Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 151 AGTCCAGTTATTGATGAGTTTACATCGCACTAAAGAGACGAACTCCGTTCACTGAT 210
Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 211 TTGTTGCCATATTGGCAGATCAAGAGCTTTATGTAAGTTTCATCCCAAGACAA 261

RESULT 9

US-10-751-235-24
; Sequence 24, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-751-235-24

Alignment Scores:
Pred. No.: 2, 73e-42 Length: 1086
Score: 335.00 Matches: 65
Percent Similarity: 90.91% Conservatives: 5
Best Local Similarity: 84.42% Mismatches: 7
Query Match: 85.24% Indels: 0
DB: 9 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-24 (1-1086)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
Db 589 CTCGAGACTTATGCTTTCAGTGGTGAACCTGTTATATATGGAAGCAGGTTTCTCAATG 648
Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
Db 649 ACATTAGATGTGATTGGTTTATCCTTGTTCACACTACAACCTTTGATTCCTCAGAT 708
Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 709 AGTCCTGTTTATGATGCTGTTTACACTGCACTCAAGAAGCTGAGGCTCGTTCTACAGAT 768
Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 769 CTTTATCACTACTGGCAGATCGATTTGCTGTGCAAGATTGTTCTCTAGACAG 819

RESULT 10

US-10-437-963-15200
; Sequence 15200, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

```
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15200
; LENGTH: 1118
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21066C.1
US-10-437-963-15200

Alignment Scores:
Pred. No.:          Length: 1118
Score:              Matches: 66
Percent Similarity: 90.91%  Conservatives: 4
Best Local Similarity: 85.71%  Mismatches: 7
Query Match:        Indels: 0
DB:                 Gaps: 0

US-10-751-235-1 (1-77) x US-10-437-963-15200 (1-1118)
QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 166 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAAATATGGAAGCAAGGTTCTCTCAATG 225
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 226 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTAACAATTTTGGATTCCTCCATCAGAT 285
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 286 AGCCCTGTATTGATGCTGTTTACACTGCACCTCAAGGAGCAAGACTTCTGTTTACAGAT 345
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 346 CTTTACCATACTGGAAGATGATTGTTGCTGTGCAAGATTGTTCTCTAGACAA 396

RESULT 11
US-10-751-235-23
; Sequence 23, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Hordeum vulgare
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1587)..(1587)
; OTHER INFORMATION: n is a, c, g, or t
US-10-751-235-23

Alignment Scores:
Pred. No.:          Length: 1638
Score:              Matches: 65
Percent Similarity: 90.91%  Conservatives: 5
Best Local Similarity: 84.42%  Mismatches: 7
Query Match:        Indels: 0
DB:                 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-23 (1-1638)
QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 166 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAAATATGGAAGCAAGGTTCTCTCAATG 225
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 226 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTAACAATTTTGGATTCCTCCATCAGAT 285
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 286 AGCCCTGTATTGATGCTGTTTACACTGCACCTCAAGGAGCAAGACTTCTGTTTACAGAT 345
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 346 CTTTACCATACTGGAAGATGATTGTTGCTGTGCAAGATTGTTCTCTAGACAA 396

US-10-751-235-1 (1-77) x US-10-751-235-22 (1-1686)
QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 622 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAAATATGGAAGCAAGGTTCTCTCAATG 681
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 682 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTAACAATTTTGGATTCCTCCATCAGAT 741
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 742 AGCCCTGTATTGATGCTGTTTACACTGCACCTCAAGGAGCAAGACTTCTGTTCTACAGAT 801
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 802 CTTTACCATACTGGAAGATGATTGTTGCTGTGCAAGATTGTTCTCTAGACAA 852

RESULT 13
US-10-751-235-27
; Sequence 27, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1587)..(1587)
; OTHER INFORMATION: n is a, c, g, or t
US-10-751-235-27

Alignment Scores:
Pred. No.:          Length: 1686
Score:              Matches: 66
Percent Similarity: 90.91%  Conservatives: 4
Best Local Similarity: 85.71%  Mismatches: 7
Query Match:        Indels: 0
DB:                 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-22 (1-1686)
QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 622 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAAATATGGAAGCAAGGTTCTCTCAATG 681
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 682 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTAACAATTTTGGATTCCTCCATCAGAT 741
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 742 AGCCCTGTATTGATGCTGTTTACACTGCACCTCAAGGAGCAAGACTTCTGTTCTACAGAT 801
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 802 CTTTACCATACTGGAAGATGATTGTTGCTGTGCAAGATTGTTCTCTAGACAA 852

RESULT 13
US-10-751-235-27
; Sequence 27, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1587)..(1587)
; OTHER INFORMATION: n is a, c, g, or t
US-10-751-235-27

Alignment Scores:
Pred. No.:          Length: 1686
Score:              Matches: 66
Percent Similarity: 90.91%  Conservatives: 4
Best Local Similarity: 85.71%  Mismatches: 7
Query Match:        Indels: 0
DB:                 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-22 (1-1686)
QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 622 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAAATATGGAAGCAAGGTTCTCTCAATG 681
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 682 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTAACAATTTTGGATTCCTCCATCAGAT 741
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 742 AGCCCTGTATTGATGCTGTTTACACTGCACCTCAAGGAGCAAGACTTCTGTTCTACAGAT 801
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 802 CTTTACCATACTGGAAGATGATTGTTGCTGTGCAAGATTGTTCTCTAGACAA 852
```

```
Db 574 CTCGAGACATATGCTTTGAGCGGTGAACCTGTTAATATGGAAGCGAGATTTCCTCAATG 633
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 634 ACACTAGATGTGATGGTTGGTTGCTTTGTTTCAACTACAACTTTGATTCCTCCATCAGAT 693
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 694 AGTCCTGTTATTGATGCTGTTTACACCGCACCTGAAAGACGAGAGGCTCGTTCTACAGAT 753
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 754 CTTTACCATACTGCGAGATTGATTGCTGTGCAAGATTGTTCTCTAGACAG 804
```

RESULT 12

```
US-10-751-235-22
; Sequence 22, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-751-235-22
```

```
Alignment Scores:
Pred. No.:          Length: 1686
Score:              Matches: 66
Percent Similarity: 90.91%  Conservatives: 4
Best Local Similarity: 85.71%  Mismatches: 7
Query Match:        Indels: 0
DB:                 Gaps: 0
```

US-10-751-235-1 (1-77) x US-10-751-235-22 (1-1686)

```
QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 622 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAAATATGGAAGCAAGGTTCTCTCAATG 681
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 682 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTAACAATTTTGGATTCCTCCATCAGAT 741
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 742 AGCCCTGTATTGATGCTGTTTACACTGCACCTCAAGGAGCAAGACTTCTGTTCTACAGAT 801
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 802 CTTTACCATACTGGAAGATGATTGTTGCTGTGCAAGATTGTTCTCTAGACAA 852
```

RESULT 13

```
US-10-751-235-27
; Sequence 27, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1587)..(1587)
; OTHER INFORMATION: n is a, c, g, or t
US-10-751-235-27
```

```
QY 1 LeuGlnProTyrAlaGluaspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20
Db 166 CTTGAGACATCTGCTTAAAGTGGCAACCTGTAAATATGGAAGCAAGGTTCTCTCAATG 225
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 226 ACTTTAGATGTGATGGTTGGTTGCTTGTTCATTAACAATTTTGGATTCCTCCATCAGAT 285
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 286 AGCCCTGTATTGATGCTGTTTACACTGCACCTCAAGGAGCAAGACTTCTGTTTACAGAT 345
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 346 CTTTACCATACTGGAAGATGATTGTTGCTGTGCAAGATTGTTCTCTAGACAA 396
```

```
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Helianthus annuus
US-10-751-235-27

Alignment Scores:
Pred. No.: 4.18e-40 Length: 624
Score: 319.00 Matches: 60
Percent Similarity: 89.61% Conservative: 9
Best Local Similarity: 77.92% Mismatches: 8
Query Match: 81.17% Indels: 0
DB: 9 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-27 (1-624)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 253 CTAAGTCATACGCGAGTGCACGCTGTTAAACATGGAGCAACAGTTTCGCAGTTA 312
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 313 ACCCTTGATGTTATCGGTCAGCCGTTATTAACATAATTTTGACTCACTTACGCCGAT 372
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 373 AGTCCTGTAATGATCTGTTATACCGCACTAAAGAAGCTGAAGCCGTTCAACTGAT 432
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 433 CTTTGGCCATATGGAAGATAAGTCGTTATGTAAGATTATACCAAGACAA 483

RESULT 14
US-10-424-599-9451
; Sequence 9451, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yinua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285694
; SEQ ID NO 9451
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108543C.1
US-10-424-599-9451

Alignment Scores:
Pred. No.: 1.19e-25 Length: 229
Score: 224.00 Matches: 46
Percent Similarity: 80.00% Conservative: 6
Best Local Similarity: 70.77% Mismatches: 13
Query Match: 57.00% Indels: 0
DB: 7 Gaps: 0

US-10-751-235-1 (1-77) x US-10-424-599-9451 (1-229)
QY 2 GlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMetThr 21
Db 29 CRAACCTGATGCACTTAATGGAAGTCTGCTCAACATCGAGGCAAAAGTTCTCGCAGCTGACT 88
QY 22 LeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAspSer 41
```

```
Db 89 CTTGATGTTACTGGATGATCTGGTCTTAATATATACCTTTGACTCGCTCAACATGACAGT 148
QY 42 ProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeu 61
Db 149 CCTGCTATTGATGTTGTATACACTGCATTTAAAGAGGCGAGGCTCGATCACTGACCT 208
QY 62 LeuProTyrTrpLys 66
Db 209 TTACCGGATTGAAA 223

RESULT 15
US-10-751-235-48
; Sequence 48, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 48
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
US-10-751-235-48

Alignment Scores:
Pred. No.: 2.05e-22 Length: 1101
Score: 210.00 Matches: 41
Percent Similarity: 72.73% Conservative: 15
Best Local Similarity: 53.25% Mismatches: 21
Query Match: 53.44% Indels: 0
DB: 9 Gaps: 0

US-10-751-235-1 (1-77) x US-10-751-235-48 (1-1101)
QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
Db 100 CTAGACAAGTATGCCGCTCAGGCACCGCTGGACATGGAAAACCTTCTTCAGCCGGCTG 159
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
Db 160 GGTCTGGACATCATCGCAAGCCGCTGTTCAACTACGACTTCGACTCGCTGGCGCACGAC 219
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
Db 220 GACCCCGTCATCCAGCGCGTGACACGTTGCTGGCGAAGCGGAGCACCCTCCACAGCG 279
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
Db 280 CCCATCGCCTACTGGAAACATTCGCCGCACTCCAGTTTGTGTGTCGCCGGGCAG 330

Search completed: January 15, 2006, 15:25:16
Job time : 619 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame plus p2n moden

Run on: January 15, 2006, 05:58:40 ; Search time 3652 Seconds
(without alignments)
1198.507 Million cell updates/sec

Title: US-10-751-235-1

Perfect score: 393

Sequence: 1 LQPYAEDGSVNMKAFSQM.....STDLLPYWKIDALCKIVPRQ 77

Scoring table:

	Xgapop 10.0	Xgapext 0.5
-Q	10.0	0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2 1/USPTO spool/US10751235/runat 12012006 120915 2519/app query.fasta 1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751235 @CNC 1 1 4939 @runat 12012006 120915 2519 -NCFU=6 -ICPU=3
-NO MNAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.in.*

3: gb.env.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pr.*

9: gb.ro.*

10: gb.scs.*

11: gb.sy.*

12: gb.un.*

13: gb.vi.*

14: gb.htg.*

15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	100.0	1620	15	AY424805 Arabidops
2	393	100.0	1858	15	AY091083 Arabidops
3	393	100.0	3304	15	AF367289 Arabidops

4	346	88.0	92611	15	ATT4D2	AL132958 Arabidops
5	337	85.8	1867	15	BT012891	BT012891 Lycopersi
6	335	85.2	1876	15	AK065689	AK065689 Oryza sat
7	303	77.1	110000	15	AP008216	Continuation (209
8	303	77.1	178022	15	AC025783	AC025783 Oryza sat
9	303	77.1	299735	15	AB017117	AB017117 Oryza sat
10	227	57.8	205408	14	AC155564	AC155564 Zea mays
11	220	56.0	110000	14	CR954213	Continuation (5 of
12	211	53.7	110000	14	AK954201	Continuation (9 of
13	206	52.4	4217	15	AK068163	AK068163 Oryza sat
14	206	52.4	110000	14	CR954209	Continuation (4 of
15	201	51.1	2038	15	AY601887	AY601887 Ginkgo bi
16	200	50.9	1831	15	AF022457	AF022457 Glycine m
17	199	50.6	2093	15	AK100596	AK100596 Oryza sat
18	199	50.6	2401	15	AK066680	AK066680 Oryza sat
19	198	50.4	1788	15	AY142017	AY142017 Arabidops
20	198	50.4	2017	15	AY058173	AY058173 Arabidops
21	198	50.4	2057	15	AY056446	AY056446 Arabidops
22	196	49.9	404	6	BD224432	BD224432 Materials
23	196	49.9	404	6	AR216482	AR216482 Sequence
24	196	49.9	404	6	AR432884	AR432884 Sequence
25	196	49.9	1886	15	BT002582	BT002582 Arabidops
26	196	49.9	1921	15	AY062675	AY062675 Arabidops
27	194	49.4	17857	14	AC155503	AC155503 Zea mays
28	191.5	48.7	1926	15	PSCYTP450	Z43263 P.sativum m
29	191	48.6	1077	15	AK220829	AK220829 Arabidops
30	179	45.5	103973	15	AP004028	AP004028 Oryza sat
31	179	45.5	110000	14	CR954215	Continuation (2 of
32	179	45.5	110000	15	AP008208	Continuation (352
33	179	45.5	145014	15	AP004048	AP004048 Oryza sat
34	177.5	45.2	2269	15	AF459441	AF459441 Skeletone
35	173.5	44.1	197405	15	ATCHRRIV40	AL161540 Arabidops
36	173.5	44.1	202860	15	ATPFC2	Z97337 Arabidops
37	159	40.5	378	15	ATH526463	ATH526463 Arabidops
38	152.5	38.8	110000	15	AP008208	Continuation (40 o
39	152.5	38.8	170051	15	AP005772	AP005772 Oryza sat
40	149	37.9	119420	15	AC079041	AC079041 Arabidops
41	131	33.3	109093	15	AC148994	AC148994 Medicago
42	131	33.3	120761	15	AC124218	AC124218 Medicago
43	127	32.3	75683	14	AC153160	AC153160 Glycine m
44	124	31.6	1498	15	AK062235	AK062235 Oryza sat
45	111.5	28.4	86539	15	AP004485	AP004485 Lotus cor

ALIGNMENTS

RESULT 1	AY424805	1620 bp	mRNA	linear	PLN 07-JAN-2004
LOCUS	AY424805	Arabidopsis thaliana chloroplast carotenoid epsilon-ring hydroxylase (LUT1) mRNA, complete cds; nuclear gene for chloroplast product.			
DEFINITION	AY424805	Arabidopsis thaliana (thale cress)			
ACCESSION	AY424805	GI:40218378			
VERSION	AY424805.1				
KEYWORDS		Arabidopsis thaliana (thale cress)			
SOURCE		Arabidopsis thaliana			
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi			
REFERENCE		1 (bases 1 to 1620)			
AUTHORS		Tian, L., Musetti, V., Kim, J., Magallanes-Lundback, M. and Dellapenna, D.			
TITLE		The Arabidopsis LUT1 locus encodes a member of the cytochrome P450 family that is required for carotenoid epsilon-ring hydroxylation activity			
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 101 (1), 402-407 (2004)			
PUBMED		14709673			
REFERENCE		2 (bases 1 to 1620)			
AUTHORS		Tian, L., Musetti, V., Kim, J., Magallanes-Lundback, M. and Dellapenna, D.			
TITLE		Direct Submission			
JOURNAL		Submitted (30-SEP-2003) Biochemistry and Molecular Biology,			

FEATURES	Michigan State University, East Lansing, MI 48824, USA		TITLE	Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.	
	source	Location/Qualifiers		Arabidopsis Full Length cDNA Clones	
gene	CDS	1..1620	JOURNAL REFERENCE	2 (bases 1 to 1858)	
		/organism="Arabidopsis thaliana"		Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Lam, B., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.	
misc_difference	3' UTR	/codon_start=1	JOURNAL	Direct Submission	
		/product="chloroplast carotenoid epsilon-ring hydroxylase"		Submitted (20-MAR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA	
misc_difference	misc_difference	/protein_id="AA083120.1"	COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.	
		/db_xref="GI:40218379"		The Salk, Stanford, PGECC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.	
misc_difference	misc_difference	2..248-42	ORIGIN	Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC) contributed equally to this work as PIs.	
		Pred. No.: 393.00		Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.	
misc_difference	misc_difference	Length: 1620	FEATURES	Location/Qualifiers	
		Matches: 77		1..1858	
misc_difference	misc_difference	Score: 393.00	source	/organism="Arabidopsis thaliana"	
		Percent Similarity: 100.00%		/mol_type="mRNA"	
misc_difference	misc_difference	Best Local Similarity: 100.00%	gene	/db_xref="taxon:3702"	
		Query Match: 100.00%		/chromosome="3"	
misc_difference	misc_difference	Indels: 0	CDS	/clone="RAFL09-78-C14 (R19513)"	
		Gaps: 0		/ecotype="Columbia"	
misc_difference	misc_difference	US-10-751-235-1 (1-77) x AY424805 (1-1620)	misc_difference	/note="This clone is in a modified pBluescript vector (PLC-1) as a BamHI/XhoI insert."	
				<1..1858	
misc_difference	misc_difference		misc_difference	/gene="At3g53130"	
				<1..1661	
misc_difference	misc_difference		misc_difference	/gene="At3g53130"	
				/codon_start=3	
misc_difference	misc_difference		misc_difference	/evidence=experimental	
				/product="putative cytochrome P450"	
misc_difference	misc_difference		misc_difference	/protein_id="AA013903.1"	
				/db_xref="GI:20258842"	
misc_difference	misc_difference		misc_difference	/translation="QIFLHSGFEKSSMESSLFPSSSSSYSLFTAKPTLLSPKPKF	
				TFSSIRSEKPKPKLETSKQSWSPDWLTLTLTKSLGKNDSEGIPIANAKLDDV	
misc_difference	misc_difference		misc_difference	ADLLGALFLPLYKWMNEYPTIRLAAGPRNFVIVSDPAIAKHVLNRYPKYAKGLVAE	
				VSEFLFGSGFAEAGPLWARRRVAVPSLHRRYLSVIVRPFCKAERLVEKLQPYAE	
misc_difference	misc_difference		misc_difference	DGSANNMEAKFSQMTLDVIGLSLNFNFDLSLTDSPVIEAVYTKAEALRSYDLPY	
				WKIDALCKIVPQVAKAEKAVTLIRFETVLDIAKKEIVRERGERINDEEVNDADPSI	
misc_difference	misc_difference		misc_difference	LRLFLASREEVSSQRLDDLSMLVAGHETTSVLTWTLYLLSKSSALRKAEFEVDRI	
				VLEGNHPAFEDIKELKYITRCINESRLPHPPVLRRAQVDFDILPGNTKYNVTGQDIM	
misc_difference	misc_difference		misc_difference	ISVYNIHRSSEVWEAEELPERFDIDGAIPINETNTDFKIFPFGGPRKCVGQDQFALM	
				EAIVLAVFLQRLNVELVDPQTISMTTGTATHTTNGLYMKVQSQR"	
misc_difference	misc_difference		misc_difference	1662..1858	
				/gene="At3g53130"	
misc_difference	misc_difference		misc_difference	1842	
				/gene="At3g53130"	
misc_difference	misc_difference		misc_difference	/note="compared to genomic sequence"	
				/replace="t"	

ORIGIN

Alignment Scores:
 Pred. No.: 2,61e-42 Length: 1858
 Score: 393.00 Matches: 77
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 15 Gaps: 0

US-10-751-235-1 (1-77) x AY091083 (1-1858)

QY 1 LeuGlnProTyrAlaGluSerAlaValAsnMetGluAlaLysPheSerGlnMet 20
 639 TTGACCCCTATGACAGACGAGTGTGTAATATGAGCGAAGTCTCTCAGATG 698
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 699 ACATTGATGTCATTGGTGTCTCTTTTAACTACAAATTCGATCTTTGACTACTGAT 758
 Db
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 759 AGTCCTGTCTATTGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTCTCTACTGAT 818
 Db
 QY 61 LeuLeuProTyrTrpLysIleAlaLeuLysLeuValLeuValProArgGln 77
 819 CTTCTGCCATATTGGAAGATCGATGATGTTGTGAAGATAGTCCCGAGACAG 869
 Db

RESULT 3

AF367289 3304 bp mRNA linear PLN 24-MAY-2001
 Arabidopsis thaliana AT3G53140/T4D2_70 mRNA, complete cds.

LOCUS AF367289

DEFINITION AF367289

ACCESSION AF367289.1

VERSION FLI CDNA.

KEYWORDS Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 3304)

REFERENCE Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,
 Bower, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
 Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
 Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 3304)

TITLE Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J.,
 Bower, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
 Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
 Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (30-MAR-2001) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
 Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H.,
 Cheuk, R., Kim, C.J., Koesema, E., Meyers, M.C., Tracy, S.B., Banh, J.,
 Bower, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,

Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
 Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Southwick, A.,
 Tang, C.C., Toriumi, M., Yamada, K., Yu, G., Davis, R.W., Theologis, A.,
 and Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.

FEATURES

source

1..3304

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="RAFL09-13-P07 (R12236)"

/ecotype="Columbia"

1..2037

2038..3117

/note="caffeic acid O-methyltransferase - like protein"

/codon_start=1

/product="AT3G53140/T4D2_70"

/protein_id="AAK56277.1"

/db_xref="GI:14194165"

/translation="MENSSSRNRLALMELANMISVPMISNAARVLGIADAIWNG
 GANSPLSAEILPLHLPSHTTIGDPENLQRLMLTSYGVSEHLVGSIRKYSLT
 DVGKTLVTDSSGLSYAAVYLQHQALMRAPLVHTAVVEPETEPYKANGAAAYAYQ
 GKSEEMGLMQKMSGVSPFKAILDYGFKSILVDVGGSGDCILRMILQQPYN
 VREGINFPLPEVKAPNIPGVTHVGGDFQVPSADAI FPKVWLITTTWDECKQIMK
 NCYNALPVGKLIACEPVLPKETDSHRTRALLEGDI FVMTIYRTKGKHTBEEFEL
 GLSAGFPFRPEYIDYFYTILFQK"

3118..3304

3'UTR

ORIGIN

Alignment Scores:

Pred. No.: 5,03e-42 Length: 3304
 Score: 393.00 Matches: 77
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 15 Gaps: 0

US-10-751-235-1 (1-77) x AF367289 (1-3304)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
 602 TTGCAGCCTTATGCAGAGACGGAAGTGTGTGAATATCGAAGCGAAGTCTCTCAGATG 661
 QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
 662 ACACCTTGATGTCATTGGGTGTCTCTTTTAACTACAATTCGATCTTTGACTACTGAT 721
 Db
 QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
 722 AGTCCTGTCTATTGAAGCTGTTTACACTGCTCTTAAAGAGCTGAGCTGCTTCTACTGAT 781
 Db
 QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuLysLysIleValProArgGln 77
 782 CTTCTGCCATATTGGAAGATCGATGCAATTGTGAAGATAGTCCCGAGACAG 832
 Db

RESULT 4

ATT4D2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1

REFERENCE

AUTHORS

Nyakatura, G., Fartmann, B., Dauner, D., Steier, W., Holland, R.,

ATT4D2 92611 bp DNA linear PLN 16-APR-2005
 Arabidopsis thaliana DNA chromosome 3, BAC clone T4D2.

AL132958

AL132958.1 GI:6434245

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1

REFERENCE

AUTHORS

Nyakatura, G., Fartmann, B., Dauner, D., Steier, W., Holland, R.,

Weicheelgartner,M., Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F.
and Salanoubat,M.

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 92611)
EU Arabidopsis sequencing project.
Direct Submission

Submitted (21-DEC-1999) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinried, FRG, E-mail:
lemcke@mips.mpg.de,mayer@mips.biochem.mpg.de Project
Coordinator: Marcel Salanoubat and Francis Quetier, Groupement
d'Interet Public, Centre National de Sequencage - GENOSCOPE; 2 rue
Gaston Cremieux, BP191, 91006 Evry Cedex, France;
http://www.genoscope.cns.fr

Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.

FEATURES
source

1..92611
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/ecotype="Columbia"
2081..2646
/gene="T4D2.10"
Join(2081..2434,2533..2646)
/gene="T4D2.10"
/note="similarity to BETA-GALACTOSIDASE PRECURSOR,
Lycopersicon esculentum, gb:P48980"

gene
CDS

/codon_start=1
/product="putative protein"
/protein_id="CAB6421.1"
/db_xref="GI:6630728"
/db_xref="GOA:Q9SCQ3"
/db_xref="InterPro:IPR000922"
/db_xref="UniProt/TREMBL:Q9SCQ3"
/translation="METLCRQHVFILLVLFHSLFLGLASKIDVSDARGIRIDGG
QKRFLTNPQHGKHAACNEEPDLGRLTRISCNPEGVYITKINFADYGNPTGCHP
RNGCGARATWRIKKNCLGKCHLLVTDEMFPSPKCKGAPMLAVETCTIA"
2081..2434
/gene="T4D2.10"
/number=1
2435..2532
/gene="T4D2.10"
/number=1
2533..2646
/gene="T4D2.10"
/number=2
3341..8521
/gene="T4D2.20"
Join(3341..3364,3812..3934,4065..4305,4394..4956,
5131..6090,6168..6490,6606..6884,6997..7249,7359..7529,
7645..7737,7816..7923,8048..8164,8257..8346,8438..8521)
/gene="T4D2.20"
/note="similarity to KIAA0010 protein, Homo sapiens,
gb:A38919"

exon

intron

exon

gene

CDS
Join(3341..3364,3812..3934,4065..4305,4394..4956,
5131..6090,6168..6490,6606..6884,6997..7249,7359..7529,
7645..7737,7816..7923,8048..8164,8257..8346,8438..8521)
/gene="T4D2.20"
/note="similarity to KIAA0010 protein, Homo sapiens,
gb:A38919"

/codon_start=1
/product="putative protein"
/protein_id="CAB6421.1"
/db_xref="GI:6630728"
/db_xref="GOA:Q9SCQ2"
/db_xref="InterPro:IPR000569"
/db_xref="InterPro:IPR000048"
/db_xref="UniProt/TREMBL:Q9SCQ2"
/translation="MDLNKKVSLRGASSGEISRDALLAKVSOERLSYARRANAA
SLFIQWRYSYVRKKAJIEIOEWENLLSCHSVTLTKSWSSVLRPLPFLFFVRSLSV
QHOKIQAREIHCMTCKLILESINNDQGNFCSLAVGTSSEKSTWACQTRMVSILC
SPLLTCTNYSYKRIQVIGVNAALLRILVLTDPKSWKIITNFENFEDAKTKIILIP
IGSKSGYTSVARYIKTLTKHTDERLVITSAVLAALRPFHVKQPAFVDDNQDPTNL
AVEVYSLITLIPRLVCYLPASALIRALKHSILMPSFHTILLKDKILNISSENSE
KQSTWETIPSGVIGNIISLATVSETDFMDQSPENPFVLYVHVIVTAAENLLSQ
VESVGQDIHLDEATSNETKGNYSKISFVEMLRPVCOQWHLAKLLAASGRIIVIA
DKDASTSSKSGSETGLGLLDIARLYSCLMRIFCVNPNVIGLPVLVNLMSFCPGVIVSLW

NSLESVLLPENGCTADDASHGSAKTSWNTRSPSEKCLKHLKNDSVNKNVNVNKPFGK
SPGREVECTSQPGSGQVNESTNDVMDVETLGRGPGVIGISKEYSCLLHLFCATYAHLL
LVLDLDDIQFEKQPFMLEKQORIASMLNTLVYGLLRGTGPGESRQLMDSAIRCLHLL
YERDCRPFPCASALMLSPGRTSREPIAFAARTHEVLPTSDVLTTPSGSVITITPHVF
PFERVHVREFISKOKASRMAGEVDAPGARSIEIVVRGHVVVEDGFOQLNSIGSRL
KSSITHSVFVNESGLPEAGLDYGGLSKEFLDITKAAPATEYGLFSQFTSRLLVPSP
SRNLLENGIQMIEFLGRIVGKALYEGILLDYSFHSVFIQKLGRYSFIDELSGLDPEL
YRLMYVKHYDGLKEILCLDFTVTEFFCGKMSIIEKPGGKDTSVTNENKMOYTHAMA
DYKLNROIVPNSNAFYRLGTLDISPAWLKLFNAHEFNOLLSSGNHDI DVDDLRRNTKY
TGVSDSRSRIKIFWEVMKGPESPERCLLKFKVTVSCSRAPLLGPKYQLQPTFIHKVSC
DTSUWAAIGQDQVERLPSPASTCYNTLKLPTFKRASTWREKLLYALTSNAGFELS"
3341..3364
/gene="T4D2.20"
/number=1
3365..3811
/gene="T4D2.20"
/number=1
3812..3934
/gene="T4D2.20"
/number=2
3935..4064
/gene="T4D2.20"
/number=2
4065..4305
/gene="T4D2.20"
/number=3
4306..4393
/gene="T4D2.20"
/number=3
4394..4956
/gene="T4D2.20"
/number=4
4957..5130
/gene="T4D2.20"
/number=4
5131..6090
/gene="T4D2.20"
/number=5
6091..6167
/gene="T4D2.20"
/number=5
6168..6490
/gene="T4D2.20"
/number=6
6491..6605
/gene="T4D2.20"
/number=6
6606..6884
/gene="T4D2.20"
/number=7
6885..6996
/gene="T4D2.20"
/number=7
6997..7249
/gene="T4D2.20"
/number=8
7250..7358
/gene="T4D2.20"
/number=8
7359..7529
/gene="T4D2.20"
/number=9
7530..7644
/gene="T4D2.20"
/number=9
7645..7737
/gene="T4D2.20"
/number=10
7738..7815
/gene="T4D2.20"
/number=10
7816..7923
/gene="T4D2.20"

```

/number=11
7524. .8047
/gene="T4D2.20"
/number=11
8048. .8164
/gene="T4D2.20"
/number=12
8165. .8256
/gene="T4D2.20"
/number=12
8257. .8346
/gene="T4D2.20"
/number=13
8347. .8437
/gene="T4D2.20"
/number=13
8438. .8521
/gene="T4D2.20"
/number=14
9057. .10601
/gene="T4D2.30"
complement(join(9057. .9151,9403. .9658,9810. .10043,
10124. .10254,10370. .10601))
/gene="T4D2.30"
complement(join(9057. .9151,9403. .9658,9810. .10043,
10124. .10254,10370. .10601))
/gene="T4D2.30"
/feature="strong similarity to proline-rich protein, Raphanus
sativa, PIR:S16748"
/codon_start=1
/product="putative protein"
/protein_id="CAB64213.1"
/db_xref="GI:6630730"
/db_xref="GOA:Q9SCQ1"
/db_xref="InterPro:IPR001087"
/db_xref="UniProt/TREMBL:Q9SCQ1"
/translation="MRVSGFRVLLVSCFFCKSKGVNPALIMFGDSIVDVGNNNLL
SIVKNSFLPYGRDFIDQRTGRCNGKLAIVFSAEYLGFSPPAPLFSREANENILI
GANFASNSGYDASVSPGISLSTQLSYRAYQNRVTRMIGRGNARILFSRGHIL
SAGSDIFQNTYINPLNLINLPDQADILLRSEFQNLNLYELGARRIGVISLPPMG
CLPAATILPGAGNCKSVRLNNDAMFNTKLENTTRLLMNRHSLGLRLVAFNVYQPLD
IITNPTDNGVYFWDGPHFTEAVNELLAQQLGQGISLIN"
Alignment Scores:
Pred. No.: 4,246-34 Length: 92611
Score: 346.00 Matches: 76
Percent Similarity: 71.03% Conservative: 0
Best Local Similarity: 71.03% Mismatches: 1
Query Match: 88.04% Indels: 31
DB: 15 Gaps: 1
US-10-751-235-1 (1-77) x ATT4D2 (1-92611)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
DB 17626 TTGAGCCTTATGCGAGAGCGAAGTCTGTGAATATGGAAGCAAGTTCTCTCAGATG 17685

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
DB 17686 ACACCTTGATGCTGGTGTCTCTTTTAACTACAAATTCGATCTCTTGACTACTGAT 17745

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 17746 AGTCCTGTCAATGAAGCTGTTTACACTGCTCTTAAAGAAGCTGAGCTGCTTCTACTGAT 17805

QY 61 LeuLeuProTyrTrpLysIle----- 67
DB 17806 CTTCTGCCATATGGAAGGC-AAGTTCTCTGTGTTTTTTTCTGTGTTTGTGTAATGTGTG 17864

QY 68 -----AspAlaLeu 70
DB 17865 GAACAAATTGGATCTTGTTAATTGAGAGGTTTGTGTTTCTTTCAGATCGATCGATG 17924

```

```

QY 71 CysLysIleValProArgGln 77
DB 17925 TGTAAGATAGTCCCGAGACAG 17945

RESULT 5
LOCUS BT012891
DEFINITION Lycopersicon esculentum clone 113997P, mRNA linear PLN 11-MAY-2004
ACCESSION BT012891
VERSION BT012891.1 GI:47104306
KEYWORDS FLI CDNA.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1967)
REFERENCE
AUTHORS Kirtness,E.P., Wang,W. and Vazeille,A.
TITLES Direct Submission
JOURNAL Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source
1. 1967
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="113997P"
/tissue_type="trichome"
/feature="HTDI88"

ORIGIN
Alignment Scores:
Pred. No.: 8,546-35 Length: 1967
Score: 337.00 Matches: 66
Percent Similarity: 92.21% Conservative: 5
Best Local Similarity: 85.71% Mismatches: 6
Query Match: 85.75% Indels: 0
DB: 15 Gaps: 0
US-10-751-235-1 (1-77) x BT012891 (1-1967)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAlaMetGluAlaLysPheSerGlnMet 20
DB 604 CTTTACCTGATGCAATTTCTGGCTCTGCAGTGAATATGGAGCAAGTTTCTCACTA 663

QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40
DB 664 ACACCTTGATGTTATTTGGCTTGCACCTCTCAATTCAAATTTTGATTCCTTACTACTGAC 723

QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 724 AGTCAGTATTATGATGCAAGTTTACACTGCACTTAAAGAAGCAAGTCCGTTCACTGAT 783

QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77
DB 784 TTGTTGCCATATGCGAGATCAAGCTTTATGTAAAGTTTCATCCACGACAA 834

RESULT 6
LOCUS AK065689
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J013039G17, full
insert sequence.
ACCESSION AK065689
VERSION AK065689.1 GI:32975707
KEYWORDS FLI CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
REFERENCE
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team;

```

Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shihiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ohtsuki, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nakamura, A., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayaahizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

12869764

2 (bases 1 to 1876)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayaahizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Kagawa, I., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007]

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shihiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kouda, M., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,

Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayaahizaki, Y.

Location/Qualifiers

source 1.1876

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clones="J013039G17"

ORIGIN

Alignment Scores:

Pred. No.: 15e-34 Length: 1876

Score: 335.00 Matches: 66

Percent Similarity: 90.91% Conservatve: 4

Best Local Similarity: 85.71% Mismatches: 7

Query Match: 85.24% Indels: 0

DB: 15 Gaps: 0

US-10-751-235-1 (1-77) x AK065689 (1-1876)

Qy 1 LeuGlnProTyrAlaGluAspGlySerAlaValaAsnMetGluAlaLysPheSerGlnMet 20

Db 726 CTTGAGACATCTGCTTTAAAGTGGCAACCTGTAATATGGAAGCAAGGTTCTCTCAATG 785

Qy 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrAsp 40

Db 786 ACTTATGAGATGATGGTTGCTTGTTCATTACATTTTGTATTCCTCCATCAGAT 845

Qy 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60

Db 846 AGCCTGTATTGATGCTGTTTACACTGCACCTCAAGGAAGCAAGAACTTCGTTCTACAGAT 905

Qy 61 LeuLeuProTyrTrpLysIleAspAlaLeuCysLysIleValProArgGln 77

Db 906 CTTTATCCATACATGGAAGATTGATTGCTGTGCAAGATTGTTCTCTAGACAA 956

RESULT 7

AP008216_208

WPCOMMENT

Sequence split into 227 fragments LOCUS AP008216 Accession AP008216

Fragment Name	Begin	End
AP008216_000	1	110000
AP008216_001	100001	210000
AP008216_002	200001	310000
AP008216_003	300001	410000
AP008216_004	400001	510000
AP008216_005	500001	610000
AP008216_006	600001	710000
AP008216_007	700001	810000
AP008216_008	800001	910000
AP008216_009	900001	1010000
AP008216_010	1000001	1110000
AP008216_011	1100001	1210000
AP008216_012	1200001	1310000
AP008216_013	1300001	1410000
AP008216_014	1400001	1510000
AP008216_015	1500001	1610000
AP008216_016	1600001	1710000
AP008216_017	1700001	1810000
AP008216_018	1800001	1910000
AP008216_019	1900001	2010000
AP008216_020	2000001	2110000
AP008216_021	2100001	2210000
AP008216_022	2200001	2310000
AP008216_023	2300001	2410000
AP008216_024	2400001	2510000
AP008216_025	2500001	2610000
AP008216_026	2600001	2710000
AP008216_027	2700001	2810000
AP008216_028	2800001	2910000
AP008216_029	2900001	3010000
AP008216_030	3000001	3110000

AP008216_031 3100001 3210000
AP008216_032 3200001 3310000
AP008216_033 3300001 3410000
AP008216_034 3400001 3510000
AP008216_035 3500001 3610000
AP008216_036 3600001 3710000
AP008216_037 3700001 3810000
AP008216_038 3800001 3910000
AP008216_039 3900001 4010000
AP008216_040 4000001 4110000
AP008216_041 4100001 4210000
AP008216_042 4200001 4310000
AP008216_043 4300001 4410000
AP008216_044 4400001 4510000
AP008216_045 4500001 4610000
AP008216_046 4600001 4710000
AP008216_047 4700001 4810000
AP008216_048 4800001 4910000
AP008216_049 4900001 5010000
AP008216_050 5000001 5110000
AP008216_051 5100001 5210000
AP008216_052 5200001 5310000
AP008216_053 5300001 5410000
AP008216_054 5400001 5510000
AP008216_055 5500001 5610000
AP008216_056 5600001 5710000
AP008216_057 5700001 5810000
AP008216_058 5800001 5910000
AP008216_059 5900001 6010000
AP008216_060 6000001 6110000
AP008216_061 6100001 6210000
AP008216_062 6200001 6310000
AP008216_063 6300001 6410000
AP008216_064 6400001 6510000
AP008216_065 6500001 6610000
AP008216_066 6600001 6710000
AP008216_067 6700001 6810000
AP008216_068 6800001 6910000
AP008216_069 6900001 7010000
AP008216_070 7000001 7110000
AP008216_071 7100001 7210000
AP008216_072 7200001 7310000
AP008216_073 7300001 7410000
AP008216_074 7400001 7510000
AP008216_075 7500001 7610000
AP008216_076 7600001 7710000
AP008216_077 7700001 7810000
AP008216_078 7800001 7910000
AP008216_079 7900001 8010000
AP008216_080 8000001 8110000
AP008216_081 8100001 8210000
AP008216_082 8200001 8310000
AP008216_083 8300001 8410000
AP008216_084 8400001 8510000
AP008216_085 8500001 8610000
AP008216_086 8600001 8710000
AP008216_087 8700001 8810000
AP008216_088 8800001 8910000
AP008216_089 8900001 9010000
AP008216_090 9000001 9110000
AP008216_091 9100001 9210000
AP008216_092 9200001 9310000
AP008216_093 9300001 9410000
AP008216_094 9400001 9510000
AP008216_095 9500001 9610000
AP008216_096 9600001 9710000
AP008216_097 9700001 9810000
AP008216_098 9800001 9910000
AP008216_099 9900001 10010000
AP008216_100 10000001 10110000
AP008216_101 10100001 10210000
AP008216_102 10200001 10310000
AP008216_103 10300001 10410000

AP008216_104 10400001 10510000
AP008216_105 10500001 10610000
AP008216_106 10600001 10710000
AP008216_107 10700001 10810000
AP008216_108 10800001 10910000
AP008216_109 10900001 11010000
AP008216_110 11000001 11110000
AP008216_111 11100001 11210000
AP008216_112 11200001 11310000
AP008216_113 11300001 11410000
AP008216_114 11400001 11510000
AP008216_115 11500001 11610000
AP008216_116 11600001 11710000
AP008216_117 11700001 11810000
AP008216_118 11800001 11910000
AP008216_119 11900001 12010000
AP008216_120 12000001 12110000
AP008216_121 12100001 12210000
AP008216_122 12200001 12310000
AP008216_123 12300001 12410000
AP008216_124 12400001 12510000
AP008216_125 12500001 12610000
AP008216_126 12600001 12710000
AP008216_127 12700001 12810000
AP008216_128 12800001 12910000
AP008216_129 12900001 13010000
AP008216_130 13000001 13110000
AP008216_131 13100001 13210000
AP008216_132 13200001 13310000
AP008216_133 13300001 13410000
AP008216_134 13400001 13510000
AP008216_135 13500001 13610000
AP008216_136 13600001 13710000
AP008216_137 13700001 13810000
AP008216_138 13800001 13910000
AP008216_139 13900001 14010000
AP008216_140 14000001 14110000
AP008216_141 14100001 14210000
AP008216_142 14200001 14310000
AP008216_143 14300001 14410000
AP008216_144 14400001 14510000
AP008216_145 14500001 14610000
AP008216_146 14600001 14710000
AP008216_147 14700001 14810000
AP008216_148 14800001 14910000
AP008216_149 14900001 15010000
AP008216_150 15000001 15110000
AP008216_151 15100001 15210000
AP008216_152 15200001 15310000
AP008216_153 15300001 15410000
AP008216_154 15400001 15510000
AP008216_155 15500001 15610000
AP008216_156 15600001 15710000
AP008216_157 15700001 15810000
AP008216_158 15800001 15910000
AP008216_159 15900001 16010000
AP008216_160 16000001 16110000
AP008216_161 16100001 16210000
AP008216_162 16200001 16310000
AP008216_163 16300001 16410000
AP008216_164 16400001 16510000
AP008216_165 16500001 16610000
AP008216_166 16600001 16710000
AP008216_167 16700001 16810000
AP008216_168 16800001 16910000
AP008216_169 16900001 17010000
AP008216_170 17000001 17110000
AP008216_171 17100001 17210000
AP008216_172 17200001 17310000
AP008216_173 17300001 17410000
AP008216_174 17400001 17510000
AP008216_175 17500001 17610000
AP008216_176 17600001 17710000

Qy 67 -----l-leaspalateuCyelylellevalproArg 76
Db 113584 TCTGATGAGCCCTGTTTGTCTTGATAGATTGCTGTGCAAGATGTTCTCTAGA 113643
Qy 77 Gln 77
Db 113644 CAA 113646
RESULT 9 299735 bp DNA linear PLN 06-JUN-2003
LOCUS AE017117
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 10, section 71 of 77 of the complete sequence.
ACCESSION AE017117 AE016959
VERSION AE017117.1 GI:31433358
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 299735)
REFERENCE
AUTHORS The Rice Chromosome 10 Sequencing Consortium
CONSTRM In-depth view of structure, activity, and evolution of rice chromosome 10
TITLE
JOURNAL Science 300, 1566-1569 (2003)
REFERENCE 2 (bases 1 to 299735)
AUTHORS Buell,C.R., Wing,R.A., McCombie,W.R., Messing,J. and Yuan,Q.
JOURNAL Direct Submission
TITLE Submitted (05-MAY-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
COMMENT This is the pseudomolecule for rice chromosome 10, which was constructed by resolving discrepancies between overlapping BACs, trimming the overlap regions, and linking the unique sequences to form a contiguous sequence. Genes in individual BAC clone were identified by a combination of several methods: Gene prediction programs, searches of the complete sequence against a peptide database and EST databases. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
FEATURES
source
1..299735 Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="10"
misc_feature
<1..133260
/note="Chromosome Sequence Derivation: nucleotide sequence in this region was derived from BAC clone OSUNBA0001014 (GB:AC025783)."
<431..>3685
/locus_tag="OSUNBA0001014.2"
/notes="EST D48843, BE230623 from this gene"
join(431..633,699..781,864..905,1063..1104,1442..1525,1702..1789,1913..2010,3644..3685)
/locus_tag="OSUNBA0001014.2"
join(471..633,699..781,864..905,1063..1104,1442..1525,1702..1789,1913..2010,3644..3685)
/locus_tag="OSUNBA0001014.2"
/codon_start=1
/product="unknown protein"
/protein_id="AAP54878.1"
/db_xref="GI:31433359"
/translation="MAMAAATPVLLLRSSATILLPRSSPRASACVLAPQRRRRGRWRLLRSSPPPEGVPGELMESEDSKFPVLNADDMYGPALLIGPEEGTNKVQEFLLR ELDEGLFKVHCTEEMTKQLWDAMTEQPSIEAVKIASMPRICIFSGLTGEMMWF

gene
mRNA
CDS
gene
mRNA
CDS
gene
mRNA
CDS
gene
mRNA
CDS

INSPYPTGLPEPAFAALVPNSADKILGEVTEEINGDHEMLLLVAGDFSRSCYC"
<6954..>9818
/locus_tag="OSUNBA0001014.3"
/note="similar to C 4 sterol methyl oxidase GB:AAAC49139
GI:1161339 (Saccharomyces cerevisiae); EST D15076, C97849
from this gene"
join(6954..7365,7735..8057,9074..9221,9316..9348,
9516..9818)
/locus_tag="OSUNBA0001014.3"
join(7036..7365,7735..8057,9074..9221,9316..9348,
9516..9587)
/locus_tag="OSUNBA0001014.3"
/codon_start=1
/product="putative C-4 sterol methyl oxidase"
/protein_id="AAP54879.1"
/db_xref="GI:31433360"
/translation="MLPYATAAEAEALGRAMTAESLWFRYSAGIDPYVLFWNHILF
LFVFTLAPLVALLELPAAGVFKLQPKVRLSREFFRCYRDVMRLFLVFLGPLQ
LVSPTVMVGIIHTGLPLPSIGENAAQLVFLVEDLYNLYWHLLHGWGEYKIHVS
HHEFTAPGPAAPVAHWAELVILGIPSGFVGALAPGHMITFELWILVLRQMEAIETHSG
FDFFNLTKYIPFYGAEYHDYHVGVGRSQSNFASVPTYCDLYLGTDKGYRYHKAYQ
AKMALGQTEGEKADSNGLSYAKLD"
complement(<10045..>13873)
/locus_tag="OSUNBA0001014.4"
/note="contains pfam profile PF00098 (zinc finger, CCHC
class); EST AU161993, AU057223, AU093006, AU057224 from
this gene"
complement(join(<10045..10997,11698..11793,12830..13057,
13670..13873))
/locus_tag="OSUNBA0001014.4"
complement(join(10454..10997,11698..11793,12830..12933))
/locus_tag="OSUNBA0001014.4"
/codon_start=1
/product="putative zinc finger protein"
/protein_id="AAP54880.1"
/db_xref="GI:31433361"
/translation="WSSRSPPPKDDRIITRTSYRDAPYRDRSRGPRFPNDLCNNC
KRGPHARDCPNALCHACGLPGHIAECSKCLWCNCKEFGHMANSCPNEGICRNCG
KSGHIAECSAPMLPGMRILCSNCKPGHLAAECTKACNNCKSGHLARNCPEP
VCNLVNSGHILARECPKSDAINEERGPPFPFGGYSDDVVCACRACQVGHMSRDCMAGAFM
ICHNCGGRGHWAYECPSGRLMDRFPFRRF"
<14813..>17815
/locus_tag="OSUNBA0001014.12"
/note="predicted by fgenesh"
join(<14813..15352,16307..16418,16786..16950,
17391..>17815)
/locus_tag="OSUNBA0001014.12"
join(14813..15352,16307..16418,16786..16950,17391..17815)
/locus_tag="OSUNBA0001014.12"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP54881.1"
/db_xref="GI:31433362"
/translation="NALVGCNNGEAAAPAAATWTPYCTVVAADMDFCVLSCPCBERA
LPDHADACAGTGRGGGPGVPARVYRLASVATHDRVVPVLFPAARVLVGCPDAE
LARFAACAGARAAEEALEGBVCVAMRAFAKGAARFRAVSVPLRDFRPLIDTL
RELYCTADPTPATSPPLETTVCFLYVMVKRKKVITLDSRRKRTESIDLAOLVEA
TMDFSKLTILDFDVIERLRNEFSNSEVPAIALIADPTCFHLHKFASRISVGCRI
NLPLNSQNTLSCVRCGLLELRAAQCKRHEADQSHSDAPRTCLCTVQHPPLNRSIQHER
LCGALNLMIDPEVLVTKMTCQRETQMCISMLDQFSFAFNKHVLTITSCMCIVWFL
LYYWRQSVPEVYHDTSSAIN"
-18650..>20461
/locus_tag="OSUNBA0001014.5"
/note="similar to cellulase GB:CAA48324 GI:311835
(Tropaeolum majus); EST C24984 from this gene"
join(18650..18935,19046..19146,19237..19454,20098..>20461)
/locus_tag="OSUNBA0001014.5"
join(18698..18935,19046..19146,19237..19454,20098..>20461)
/locus_tag="OSUNBA0001014.5"
/codon_start=1
/product="putative cellulase"
/protein_id="AAP54882.1"
/db_xref="GI:31433363"

Project information		Features		Location/Qualifiers	
Web site: http://www.tigr.org/tdb/tgi/maize/		source		1. .205408	
Contact: maize@tigr.org				/organism="Zea mays"	
				/mol_type="genomic DNA"	
				/strain="B73"	
				/db_xref="taxon:4577"	
				/clone="ZMMBc0171N08"	
				1703. .1802	
		gap		/estimated_length=unknown	
		gap		4663. .4762	
		gap		11548. .11647	
		gap		/estimated_length=unknown	
		gap		15559. .15658	
		gap		/estimated_length=unknown	
		gap		18073. .18172	
		gap		/estimated_length=unknown	
		gap		22716. .22815	
		gap		/estimated_length=unknown	
		gap		35447. .35546	
		gap		/estimated_length=unknown	
		gap		59876. .59975	
		gap		/estimated_length=unknown	
		gap		61762. .61861	
		gap		/estimated_length=unknown	
		gap		64114. .64213	
		gap		/estimated_length=unknown	
		gap		68253. .68352	
		gap		/estimated_length=unknown	
		gap		71636. .71735	
		gap		/estimated_length=unknown	
		gap		74352. .74451	
		gap		/estimated_length=unknown	
		gap		78945. .79044	
		gap		/estimated_length=unknown	
		gap		82989. .83088	
		gap		/estimated_length=unknown	
		gap		84258. .84357	
		gap		/estimated_length=unknown	
		gap		88369. .88468	
		gap		/estimated_length=unknown	
		gap		103421. .103520	
		gap		/estimated_length=unknown	
		gap		111901. .112000	
		gap		/estimated_length=unknown	
		gap		139515. .139614	
		gap		/estimated_length=unknown	
		gap		141413. .141512	
		gap		/estimated_length=unknown	
		gap		146354. .146453	
		gap		/estimated_length=unknown	
		gap		150399. .150498	
		gap		/estimated_length=unknown	
		gap		152461. .152560	
		gap		/estimated_length=unknown	
		gap		153780. .153879	
		gap		/estimated_length=unknown	
		gap		167813. .167912	
		gap		/estimated_length=unknown	
		gap		170603. .170702	
		gap		/estimated_length=unknown	
		gap		198609. .198708	
		gap		/estimated_length=unknown	
		gap		201018. .201117	
		gap		/estimated_length=unknown	
		gap		204280. .204379	
		gap		/estimated_length=unknown	
		ORIGIN			
		Alignment Scores:		8.48e-18	
		Pred. No.:		227.00	
		Score:		80.88%	
		Percent Similarity:			
				Length:	
				Matches:	
				Conservative:	
				205408	
				49	
				6	

* NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1702: contig of 1702 bp in length

* 1703 1802: gap of unknown length

* 1803 4662: contig of 2860 bp in length

* 4663 4762: gap of unknown length

* 4763 11547: contig of 6785 bp in length

* 11548 11647: gap of unknown length

* 11648 15558: contig of 3911 bp in length

* 15559 15658: gap of unknown length

* 15659 18072: contig of 2414 bp in length

* 18073 18172: gap of unknown length

* 18173 22715: contig of 4543 bp in length

* 22716 22815: gap of unknown length

* 22816 35446: contig of 12631 bp in length

* 35447 35546: gap of unknown length

* 35547 59875: contig of 24329 bp in length

* 59876 61761: contig of 1786 bp in length

* 61762 61861: gap of unknown length

* 61862 64113: contig of 2252 bp in length

* 64114 68252: gap of unknown length

* 68253 68352: contig of 4039 bp in length

* 68353 71635: contig of 3283 bp in length

* 71636 71735: gap of unknown length

* 71736 74351: contig of 2616 bp in length

* 74352 74451: gap of unknown length

* 74452 78944: contig of 4493 bp in length

* 78945 79044: gap of unknown length

* 79045 82988: contig of 3944 bp in length

* 82989 83088: gap of unknown length

* 83089 84257: contig of 1169 bp in length

* 84258 84357: gap of unknown length

* 84358 88368: contig of 4011 bp in length

* 88369 88468: gap of unknown length

* 88469 103420: contig of 14952 bp in length

* 103421 103520: gap of unknown length

* 103521 111900: contig of 8380 bp in length

* 111901 112000: gap of unknown length

* 112001 139514: contig of 27514 bp in length

* 139515 139614: gap of unknown length

* 139615 141412: contig of 1798 bp in length

* 141413 141512: gap of unknown length

* 141513 146353: contig of 4841 bp in length

* 146354 146453: gap of unknown length

* 146454 150398: contig of 3945 bp in length

* 150399 150498: gap of unknown length

* 150499 152460: contig of 1962 bp in length

* 152461 152560: gap of unknown length

* 152561 153779: contig of 1219 bp in length

* 153780 153879: gap of unknown length

* 153880 167812: contig of 13933 bp in length

* 167813 167912: gap of unknown length

* 167913 170602: contig of 2690 bp in length

* 170603 170702: gap of unknown length

* 170703 198608: contig of 27906 bp in length

* 198609 198708: gap of unknown length

* 198709 201017: contig of 2309 bp in length

* 201018 201117: gap of unknown length

* 201118 204279: contig of 3162 bp in length

* 204280 204379: gap of unknown length

* 204380 205408: contig of 1029 bp in length.

Best Local Similarity: 72.06% Mismatches: 12
Query Match: 57.76% Indels: 1
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x AC155564 (1-205408)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
DB 108500 CTTGAGCCATATGCTTTGGTGGGAACTGTCAATATAGAAAGGATTTCTCAGTTG 108559
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 108560 ACATTGGATGTTATGGTTGTCATTGTTCACTACACTTTTATCCCTCAGATAGAC 108619
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 108620 AGTTCTAGCATGTTGTTGTTATATCTGCACTCAAGAAGCAGAGATTGTTCTACAGAT 108679
QY 61 Leu-LeuProTyrTrpLysIle 67
DB 108680 CTTTTCCTACTAGAGGTT 108701

RESULT 11
CR954213_4
WPCOMMENT

Sequence split into 6 fragments LOCUS CR954213 Accession CR954213

Fragment Name	Begin	End
CR954213_0	1	110000
CR954213_1	100001	210000
CR954213_2	200001	310000
CR954213_3	300001	410000
CR954213_4	400001	510000
CR954213_5	500001	531849

Continuation (5 of 6) of CR954213 from base 400001 (CR954213 Ostreococcus tauri strain Q

Alignment Scores:
Pred. No.: 3.6e-17 Length: 110000
Score: 220.00 Matches: 41
Percent Similarity: 75.71% Conservative: 12
Best Local Similarity: 58.57% Mismatches: 17
Query Match: 55.98% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x CR954213_4 (1-110000)

QY 8 GlySerAlaValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIleGlyLeu 27
DB 11563 GGTGAATCGTGGAGATGGAGAACTTTTACTCTCGTTTGGCTGGATATCATCGTAAG 11622
QY 28 SerLeuPheAsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGluAlaVal 47
DB 11623 GCTGTGTTCACTACGATTTTCGATTTCTTTGACACGAGCATCCGCTGATCAAGGCTGTG 11682
QY 48 TyrThrAlaLeuLysGluAlaGluLeuArgSerThrAspLeuLeuProTyrTrpLysIle 67
DB 11683 TACACGCTTCGCGGAGCGGAGATCGAAGCGTGACGTTTCATTCCTGATTTGGAAGGTG 11742
QY 68 AspAlaLeuLysIleValProArgGln 77
DB 11743 CCGCGCTTCGCTGCTGCTGCTCGGACAG 11772

RESULT 12
CR954201_08/c
WPCOMMENT

Sequence split into 11 fragments LOCUS CR954201 Accession CR954201

Fragment Name	Begin	End
CR954201_00	1	110000
CR954201_01	100001	210000
CR954201_02	200001	310000
CR954201_03	300001	410000
CR954201_04	400001	510000
CR954201_05	500001	610000
CR954201_06	600001	710000

CR954201_07 700001 810000
CR954201_08 800001 910000
CR954201_09 900001 1010000
CR954201_10 1000001 1076297
Continuation (9 of 11) of CR954201 from base 800001 (CR954201 Ostreococcus tauri strain

Alignment Scores:
Pred. No.: 5.74e-16 Length: 110000
Score: 211.00 Matches: 41
Percent Similarity: 68.83% Conservative: 12
Best Local Similarity: 53.25% Mismatches: 24
Query Match: 53.69% Indels: 0
DB: 14 Gaps: 0

US-10-751-235-1 (1-77) x CR954201_08 (1-110000)

QY 1 LeuGlnProTyrAlaGluAspGlySerAlaValAsnMetGluAlaLysPheSerGlnMet 20
DB 85758 CTCGACAAATCGCGGAGGATGGAGAACTGTAGACATGGAGCGGTTCTGCTCGGTG 85699
QY 21 ThrLeuAspValIleGlyLeuSerLeuPheAsnTyrAsnPheAspSerLeuThrThrAsp 40
DB 85698 AGTCTGGACATCATCGCTTGGCTTCACTACGACTTTGGTTCTAGTCAGCAAGGAG 85639
QY 41 SerProValIleGluAlaValTyrThrAlaLeuLysGluAlaGluLeuArgSerThrAsp 60
DB 85638 TCTCAATCATCAGCGCGTGTATTAACCTTCAGGAAGCGGCATCGAAGTACGTTTC 85579
QY 61 LeuLeuProTyrTrpLysIleAspAlaLeuLysIleValProArgGln 77
DB 85578 TATTTCCTATTTGGACATTCGCTTCGCGAGGACATCGTCCCTCGCCAG 85528

RESULT 13
AK068163
LOCUS

DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:J013135B04, full insert sequence.
AK068163
VERSION
AK068163.1 GI:32978181
KEYWORDS
FLI CDNA; CAP trapper.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
1

The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team:
Kituchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, B., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kogawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oda, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)

TITLE

JOURNAL
PUBMED
REFERENCE
2 (bases 1 to 4217)
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayaehizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,

Seoul 151-742, Republic of Korea
Location/Qualifiers
1. .2038
/organism="Ginkgo biloba"
/mol_type="mRNA"
/db_xref="taxon:3311"
5'UTR
CDS
11..1771
/codon_start=1
/product="putative 97B2-like cytochrome P450"
/protein_id="AAT28222.1"
/db_xref="GI:47498772"
/translation="MITVPSIPCSSEKDTVQPIGRKVLTVENVVSCGVSSSFLPGS
KFPFCESSSTERATKSNRTLLDNASNFTNLISGQLGTMETIARGAVSDLFQKELFF
SYDPIETGHSYVKLAFGPKSFVVVSDPIVVRHILKRNAPAYDKGVLADILEPIMKKG
LIPADIGTWKPKRAIVPGFAAFLEAMVKVFGDCSERTINKLQSLLLAAEADKTMHI
DYDMEAFENALDIIIGSVFNFGSVTRSPVKAIVGTILFEAHRSTFYIPYWK
PLARLVPRQRFHEDLKIINECLDSLIQAKETQEDDIEALQGRDYSKVDASLLR
FLVDKGVVDVNGQLRDDLMWLIAGHETTAAVLTAFLAONTDKLVKAQAEIDTI
LQQRKTFEDIKRLQYIRLIIAALRLYQPPELLIRRLRQDTIFGGYRGDKGGHLP
KGTDFISVYNLHRSFYWDQPHFRPERFLIPKONKGIEGWAGFDPYRSQALYPNE
VLADFAFLPFGGPRKCVGDQFALMESTVGLAMLLQKEVELKDPPESLRIIVTGATIH
TQNLWCKIKRRKHYSESLQ"

ORIGIN

Alignment Scores:
Pred. No.: 1.35e-16 Length: 2038
Score: 201.00 Matches: 40
Percent Similarity: 76.12% Conservative: 11
Best Local Similarity: 59.70% Mismatches: 16
Query Match: 51.15% Indels: 0
DB: 15 Gaps: 0

US-10-751-235-1 (1-77) x AY601887 (1-2038)

Qy 11 ValAsnMetGluAlaLysPheSerGlnMetThrLeuAspValIleGlyLeuSerLeuPhe 30
Db 668 GYGACATGGAGCAGCAATTTCAATCTAGCTCTTGACATTATTGGGCTAAGTGCTTC 727
Qy 31 AsnTyrAsnPheAspSerLeuThrThrAspSerProValIleGluAlaValTyrThrAla 50
Db 728 AATTAGACTTTGGATCTGTACTAGAGAAATCACCAGTAATTAAGGCAGCTCTATGGTACG 787
Qy 51 LeuLysGluAlaGluLeuArgSerThrAspLeuLeuProTyrTyrIleAspAlaLeu 70
Db 788 TTGTTTGAAGCTGAGCAGACATCACTTTTACATTCCATCTGCAAAATTCCTCTTGCA 847
Qy 71 CysLysIleValProArgGln 77
Db 848 AGATGGCTAGTTCCTCGACAA 868

Search completed: January 15, 2006, 14:22:26
Job time : 3726 secs

This Page Blank (uspto)

GenCore version 5.1.1.6

Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 20:19:34 ; Search time 12340 Seconds
(without alignments)
11364.084 Million cell updates/sec

Title: US-10-751-235-5

Perfect score: 2467

Sequence: 1 atggagcttcactcttttc.....tgaagggtgagccaaaggtaa 2467

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sv.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2467	100.0	92611	15	ATT4D2
2	803	32.5	1620	15	AY424805
3	803	32.5	1858	15	AY091083
4	803	32.5	3304	15	AF367289
5	376.4	15.3	378	15	ATH526463
6	268.6	10.9	1077	15	AK220829
7	252	10.2	1967	15	BT012891
8	239.6	9.7	110000	15	AF008216_208
9	239.6	9.7	178022	15	AC025783
10	239.6	9.7	299735	15	AE017117
11	183	7.4	1876	15	AK065689
12	146.8	6.0	205408	14	AC155564
13	91	3.7	1788	15	AY142017
14	91	3.7	2017	15	AY058173
15	91	3.7	2057	15	AY056446
16	83.4	3.4	4217	15	AK068163
17	75.2	3.0	178657	14	AC155503
18	74.8	3.0	110000	14	CR954209_3

19	73.2	3.0	1831	15	AF022457
20	72.8	3.0	1926	15	PSCYTP450
21	72	2.9	1886	15	BT002582
22	72	2.9	1921	15	AY062675
23	71.8	2.9	110000	14	CR954213_4
24	71.4	2.9	110000	14	CR954201_08
25	71.2	2.9	625	15	AK221212
26	70.4	2.9	119420	15	AC079041
27	67	2.7	2093	15	AK100596
28	67	2.7	2401	15	AK066680
29	66.4	2.7	181167	14	AC155486
30	66.2	2.7	2038	15	AY601887
31	65.2	2.6	103973	15	AP004028
32	65.2	2.6	110000	15	AP008208_351
33	65.2	2.6	145014	15	AP004048
34	59.4	2.4	197405	15	ATCHRIV40
35	59.4	2.4	202860	15	ATFCA2
36	59	2.4	109093	15	AC148994
37	59	2.4	120761	15	AC124218
38	58.2	2.4	110000	14	CR954215_1
39	57.8	2.3	404	6	BD224432
40	57.8	2.3	404	6	AR216482
41	57.8	2.3	404	6	AR432884
42	56.6	2.3	2269	15	AF459441
43	55.2	2.2	207718	14	AC155487
44	53.8	2.2	86539	15	AP004485
45	53	2.1	110000	1	CP000001_30

ALIGNMENTS

RESULT 1

ATT4D2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

```
/note="similarity to BETA-GALACTOSIDASE PRECURSOR,
Lycopodium esculentum, gb:P48980"
/codon_start=1
/product="putative protein"
/protein_id="CAB64211.1"
/db_xref="GI:6630728"
/db_xref="GOA:Q9SCQ3"
/db_xref="InterPro:IPR000922"
/db_xref="UniProt/TREMBL:Q9SCQ3"
/translation="METLQCRHQHVIFILLVLFHSSLFGLASKIDVSDDARGIRIDGG
QKRFLTSPQHGKHAACNEEPDLGLTRISCNPEGYVITKINPADYGNPTGCGHF
RRGNGCARATWRIVKKNCLGKCHLLVTDEMFGPSCKGAPMLAVETCTIA"
2081..2434
/feature="T4D2.10"
/number=1
2435..2532
/feature="T4D2.10"
/number=1
2533..2646
/feature="T4D2.10"
/number=2
3341..8521
/feature="T4D2.20"
/number=7
join(3341..3364,3812..3934,4065..4305,4394..4956,
5131..6090,6168..6490,6606..6884,6997..7249,7359..7529,
7645..7737,7816..7923,8048..8164,8257..8346,8438..8521)
/feature="T4D2.20"
/number=8
/feature="similarity to KIAA0010 protein, Homo sapiens,
gb:A38919"
/codon_start=1
/product="putative protein"
/protein_id="CAB64212.1"
/db_xref="GI:6630729"
/db_xref="GOA:Q9SCQ2"
/db_xref="InterPro:IPR000048"
/db_xref="InterPro:IPR000569"
/db_xref="UniProt/TREMBL:Q9SCQ2"
/translation="MDLNRKHVSLRGASSGEISRDLALKVSOERELRSVARRANAA
SLFIQVRVKAAIEIOEWENLLSCTVLTKSWSSRVLRFLFPRVRLSV
QHQTQARIEHCMTQCKILLISINSDQGNFCSLAVGTSEDSKIWACQTRRWSVLC
SFLTCKYQSRIRKIDVGNALLRIILVLPKSWKIITNENFEDAKKIIQF
IGSCSGYYSARRIKTLTKTDERLVITTSATLALRPHVQKPAFVDNQPTNL
AVEEVSLITLPLVCVLPALIRALKHKSILMPSTLILKDKILNIISEMENSE
KQSCWTEIPVGVWIGNIISLATVSETDPMQESNPEMFYLVYVHVIWLAENLSQ
VESVGIDILHDIETASNETKGNYSKISFVEMLRPVCCQHLAKLLAASKEIRVIA
DKDASTSKGSETLGLLDIARLYSCMLRIFCVMPVLGPPVPLNLLSCPGYIIVSLW
NSLESVLLPENGCTADASHGSAKTSWNTSPSEKKLKHKNDSYKNWVNLKFSKG
SPGRHEVCTSDQPGSQVNESTNDVMDVETLRGPGVIGISKEVSCILLHLFCATYAH
LVVLDDIOPYEKQVPFMLEKQORIASMLNTLVYGLLGTGPESQLMDSARCLHLL
YRDCRHCFCSALMLSPTGTSRPTIAPAAATHEVLPSTSDVLTTPSMGSVITITPHV
PPEERHVHREFISKDKARKWAGEVDAPGARSIEIVNRGHVVEDGQQQLNSIGSRL
KSSIHVSFVNESGLPEAGLDYGLSKFELDITKAAPATEYGLGFSQTPTSRLLVPSP
SARHLENGIOMTEFLGRIVGKALYEGILLDYSFVFIQKLGYSFIDEISGLDPEL
YNNLVVHYDGLKELCLDFTVTEFCFGKMSIIEIKPGCKDTSVTNENKMYIHAMA
DYKLRQIVPFSNAFVGTJDLISPAWLKLFNAHEPNOLLGSHNDIDVDDLRNRTKY
TCGISDSSRTIKIFWEWMKGFPSEKCLLKKFTVSCSRAPLLGFKYLOPTFIHKVSC
DTSLWAAIGQGVRLPSASTCYNTLKLPTYKRASTMREKLLIATISNAGFELS"
3341..3364
/feature="T4D2.20"
/number=1
3365..3811
/feature="T4D2.20"
/number=1
3812..3934
/feature="T4D2.20"
/number=2
3935..4064
/feature="T4D2.20"
/number=2
4065..4305
/feature="T4D2.20"
/number=3
4306..4393
/feature="T4D2.30"
/number=3
complement(join(9057..9151,9403..9658,9810..10043,
10124..10254,10370..10601))
/feature="T4D2.30"
/number=3
complement(join(9057..9151,9403..9658,9810..10043,
10124..10254,10370..10601))
/feature="T4D2.30"
/number=3
```


QY 661 CGGTGGTTCCATCGCTTTCACAGAGGTATTTGCTGTGATTTGTGGAGAGTATTTCTGCA 720
Db 512 CGGTGGTTCCATCGCTTTCACAGAGGTATTTGCTGTGATTTGTGGAGAGTATTTCTGCA 571
QY 721 AATGTGCAGAGAGGCTTGTGAGAGTGTGAGGCTTTATGCAGAGACGGAAGTGTGTGCA 780
Db 572 AATGTGCAGAGAGGCTTGTGAGAGTGTGAGGCTTTATGCAGAGACGGAAGTGTGTGCA 631
QY 781 ATATGGAAGCGAAGTCTCTCAGATGACACCTTGATGATCATTTGGTGTCTCTTTTAACT 840
Db 632 ATATGGAAGCGAAGTCTCTCAGATGACACCTTGATGATCATTTGGTGTCTCTTTTAACT 691
QY 841 ACAATTTCCATCTTTGACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 692 ACAATTTCCATCTTTGACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 751
QY 901 AAGAAGCTGAGCTTCT 960
Db 752 AAGAAGCTGAGCTTCT 793
QY 961 TTTTCTGTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 794 ----- 793
QY 1021 TTGTTTTTTTTCAGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 794 -----AGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
QY 1081 GGCTGTAACTTTGATAAGGGAACCTGTTGAAGACCTTTATGCTGAAGTGTAAAGAAATGT 1140
Db 843 GGCTGTAACTTTGATAAGGGAACCTGTTGAAGACCTTTATGCTGAAGTGTAAAGAAATGT 902
QY 1141 CGAAGAGAGCGGAAGAAATCAATGATGAGGAGTATGTAATGATGCTGACCCCAAGTAT 1200
Db 903 CGAAGAGAGCGGAAGAAATCAATGATGAGGAGTATGTAATGATGCTGACCCCAAGTAT 962
QY 1201 CCTGCGTTTCTGCTTGCAGCAGAGAGAGGTTTAACTTTTCTTAACTTTTCTTAACTTATAG 1260
Db 963 CCTGCGTTTCTGCTTGCAGCAGAGAGG----- 991
QY 1261 CAAATTTGGCCCTTCATTATCGATAATCGAAGCTGATGTTGCTGAGGCTTTTTCAG 1320
Db 992 -----AG 993
QY 1321 GTATCAAGTGTGAGTACGGGATGATCTTCTCAATGCTGCTAGCGGGTCAATGAACC 1380
Db 994 GTATCAAGTGTGAGTACGGGATGATCTTCTCAATGCTGCTAGCGGGTCAATGAACC 1053
QY 1381 ACTGATCTGCTCCTCACTTGGACACTTTATCTCCTAAGTAAAGGTACCTTAAATGATCTTC 1440
Db 1054 ACTGATCTGCTCCTCACTTGGACACTTTATCTCCTAAGT----- 1093
QY 1441 TACTTTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTCTGTTCTCATTTACCTCTT 1500
Db 1094 ----- 1093
QY 1501 CAAATTTCTATGTTTATAGAACTCATCTGCAATTAAGGAAAGCAACAAGAGAGTAGACA 1560
Db 1094 -----AGAACTCATCTGCAATTAAGGAAAGCAACAAGAGAGTAGACA 1135
QY 1561 GAGTGTTAAGAGGAAGAACCCGGCTTTCAGGATATTAAGAGGTTGAAGTACATCATC 1620
Db 1136 GAGTGTTAAGAGGAAGAACCCGGCTTTCAGGATATTAAGAGGTTGAAGTACATCATC 1195
QY 1621 GTTGTATAAAGAGTCAATGCTCTATCTCCTCATCTCCTGTAAGCAATCAAGCTCATC 1680
Db 1196 GTTGTATAAAGAGTCAATGCTCTATCTCCTCATCTCCTGTAAGCAATCAAGCTCATC 1236
QY 1681 TCTCTAATTAATTCATGAACATAAATTTTCTGATGATGATGATGATGATGATGATGATGAT 1740
Db 1237 -----GTCTTGATAA 1246
QY 1741 GAAGAGCTCAAGTTCCTGACATTTCTTCTGGAAGCTATAAGGTCAATACCGGACAAGACA 1800

Db 1247 GAAGAGCTCAAGTTCCTGACATTTCTTCTGGAACCTATTAAGGTCAATACCGGACAAGACA 1306
QY 1801 TTATGATTTTCACTCTATTAACATCCATCGTTCTTCTTCCGAGGTA 1841
Db 1307 TTATGATTTTCACTCTATTAACATCCATCGTTCTTCTTCCGAGGTA 1347

AY091083 1858 bp mRNA linear PLN 18-SEP-2002
Arabidopsis thaliana putative cytochrome P450 (At3g53130) mRNA,
partial cdb.
ACCESSION AY091083
VERSION AY091083.1 GI:20258841
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1858)
AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 1858)
AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (20-MAR-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAPL cDNAs: Yamada,K., Banh,J.,
Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGSC) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGSC)
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
FEATURES
source
1. .1858
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAPL09-78-C14 (R19513)"
/ecotype="Columbia"


```
QY 781 ATATGGAAGCGAAGTCTCTCAGATGACACATTTGATGCTATTTGGGTGCTCTTTTAACT 840
Db 636 ATATGGAAGCGAAGTCTCTCAGATGACACATTTGATGCTATTTGGGTGCTCTTTTAACT 695
QY 841 ACAATTTCCGATCTTTGACTACTGATAGTCTCTGTCATTTGAAGCTGTTTACACTGCTCTTA 900
Db 696 ACAATTTCCGATCTTTGACTACTGATAGTCTCTGTCATTTGAAGCTGTTTACACTGCTCTTA 755
QY 901 AAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGAAGCAAGTTTCTGCTGTT 960
Db 756 AAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGA----- 797
QY 961 TTTTCTGCTGTTGTTGATTTGTTGGCAACAATTGGAATCTTGTAAATTGAGAGGGTTGG 1020
Db 798 ----- 797
QY 1021 TTGTTTTTTTTCAGATCGATGCATTTGTTAAGATAGTCCCGAGACAGGTGAAAGCTGAAAA 1080
Db 798 -----AGATCGATGCATTTGTTAAGATAGTCCCGAGACAGGTGAAAGCTGAAAA 846
QY 1081 GCCTGTAACTTTGATAAGGAAACTGTTTGAAGACCTTATTTGCTAAGTGTAAAGAAATTGT 1140
Db 847 GCCTGTAACTTTGATAAGGAAACTGTTTGAAGACCTTATTTGCTAAGTGTAAAGAAATTGT 906
QY 1141 CGAAAGAGAGCGCGAAGCAATCAATGATGAGGAGTATCTTAATGATGCTGACCCCAAGTAT 1200
Db 907 CGAAAGAGAGCGCGAAGCAATCAATGATGAGGAGTATCTTAATGATGCTGACCCCAAGTAT 966
QY 1201 CCTGCGTTTCTTCTTGTGCAAGCAGAGAGAGGTTTAAACTTTTTCCTTAAGTTTATAAG 1260
Db 967 CCTGCGTTTCTTCTTGTGCAAGCAGAGAGG----- 995
QY 1261 CAAATTTGGCCTTTCAATTATCGCATATATCGAAGCTGATGTTGCTATGAGGGTTTTCAG 1320
Db 996 -----AG 997
QY 1321 GTATCAAGTGTGCGATTACGGGATGATCTTCTCTCAATGCTCGTAGCGGGTCATGAACC 1380
Db 998 GTATCAAGTGTGCGATTACGGGATGATCTTCTCTCAATGCTCGTAGCGGGTCATGAACC 1057
QY 1381 ACTGATCTGCTCCTCACTTGGACACTTTATCTCTCAAGTAAGGTACCTTAAATGTATCTTC 1440
Db 1058 ACTGATCTGCTCCTCACTTGGACACTTTATCTCTCAAGTA----- 1097
QY 1441 TACTTTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTCTGTTCTCAATTTACCTCTT 1500
Db 1098 ----- 1097
QY 1501 CAAATTTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAGCAACAAGAGTAGACA 1560
Db 1098 -----AGAACTCATCTGCAATTAAGGAAGCAACAAGAGTAGACA 1139
QY 1561 GAGTGTTAGAAGGAAGAAACCGGCTTTTCAGGATATAAAGAGTTGAAGTACATCACTC 1620
Db 1140 GAGTGTTAGAAGGAAGAAACCGGCTTTTCAGGATATAAAGAGTTGAAGTACATCACTC 1199
QY 1621 GTTGTATAAACGAGTCAATGCGCTCTCATCTCATCTCTCTGTAAGCAATCAAGCTCATC 1680
Db 1200 GTTGTATAAACGAGTCAATGCGCTCTCATCTCATCTCTCT----- 1240
QY 1681 TCTTAATTTATCATGAACATAAATTTCTGATGATTTGTTCTCTGTTAGTCTTTGATAA 1740
Db 1241 -----GTCCTTGATAA 1250
QY 1741 GAAGAGCTCAAGTTCTGACATTTCTTCTGGGAACATATAAGGTCAATACCGGACAGACA 1800
Db 1251 GAAGAGCTCAAGTTCTGACATTTCTTCTGGGAACATATAAGGTCAATACCGGACAGACA 1310
QY 1801 TTATGATTTTCAGTCTATAACATCCATCGTTCTTCCGAGGTA 1841
Db 1311 TTATGATTTTCAGTCTATAACATCCATCGTTCTTCCGAGGTA 1351
```

```
RESULT 5
ATHS26463 378 bp DNA linear PLN 29-MAR-2003
LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone
DEFINITION
ACCESSION AJ526463
VERSION AJ526463.1 GI:26794723
KEYWORDS left border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
Chauvin, S., Bechtold, N., Cruaud, C., Derose, R., Pelletier, G.,
Lepiniec, B., Caboche, M. and Lecharny, A.
T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
EMBO Rep. 3 (12), 1152-1157 (2002)
12446565
2 (bases 1 to 378)
Balzergue, S.
Direct Submission
Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (http://www.genoplante.com and
http://genoplante-info.infobiogen.fr).
FEATURES
Location/Qualifiers
1..378
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="120811"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Wassilewskija"
misc_feature
1..378
/note="T-DNA flanking sequence
left border"
ORIGIN
Query Match 15.3%; Score 376.4; DB 15; Length 378;
Best Local Similarity 99.7%; Pred. No. 2.4e-79; Mismatches 1; Indels 0; Gaps 0;
Matches 377; Conservative 0;
QY 858 ACTACTGATAGTCCCTGTCATTGCAAGCTGTTTACACTGCTCTTAAAGAAAGCTGAGCTTCGT 917
Db 1 ACTACTGATAGTCCCTGTCATTGCAAGCTGTTTACACTGCTCTTAAAGAAAGCTGAGCTTCGT 60
QY 918 TCTACTGATCTTCTGCGCATATTTGGAAGCAAGTTCTCTGTTGTTTCTGCTGTTGTTG 977
Db 61 TCTACTGATCTTCTGCGCATATTTGGAAGCAAGTTCTCTGTTGTTTCTGCTGTTGTTG 120
QY 978 ATTGTTGCAACAATTTGGATTTCTTTAATTCAGAGGGTTTGGTCTGTTTTTTCAGATCG 1037
Db 121 ATTGTTGGAACAATTTGGATTTCTTTAATTCAGAGGGTTTGGTCTGTTTTTTCAGATCG 180
QY 1038 ATGCAATTGTGTAAAGTAGTCCCGAGACAGGTGAAAGCTGAAAGCGCTGTAACCTTTGATAA 1097
Db 181 ATGCAATTGTGTAAAGTAGTCCCGAGACAGGTGAAAGCTGAAAGCGCTGTAACCTTTGATAA 240
QY 1098 GGGAAACTGTTGAAGACCTTATTTGCTAAGTGTAAAGAAATTTGTCGAAGAGGCGGAAA 1157
Db 241 GGGAAACTGTTGAAGACCTTATTTGCTAAGTGTAAAGAAATTTGTCGAAGAGGCGGAAA 300
```


QY 1158 GAATCAATGATGAGGAGTATGTAATGATGCTGACCCAAAGTATCTGCGCTTTCTTGCTTG 1217
 |||||
 Db 301 GAATCAATGATGAAGAGTATGTAATGATGCTGACCCAAAGTATCTGCGCTTTCTTGCTTG 360
 |||||
 QY 1218 CAAGCAGAGAGAGGTTTT 1235
 |||||
 Db 361 CAAGCAGAGAGAGGTTTT 378
 |||||

RESULT 6
 AK220829
 LOCUS Arabidopsis thaliana gene for Cytochrom P450 -like protein, partial cds, clone: RAFL22-34-N18.
 DEFINITION
 AK220829
 ACCESSION
 AK220829.1 GI:62319016
 VERSION
 FLI CDNA.
 KEYWORDS
 Arabidopsis thaliana (thale cress)
 SOURCE
 Arabidopsis thaliana
 ORGANISM
 Arabidopsis thaliana

REFERENCE
 1 Totoki Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K., Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K., Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs
 Unpublished

TITLE
 JOURNAL
 REFERENCE
 2 (bases 1 to 1077)
 AUTHORS
 Totoki, Y., Seki, M., Ishida, J., Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Shin-i, T., Nakagawa, M., Sakamoto, N., Oishi, K., Kohara, Y., Kobayashi, M., Toyoda, A., Sakaki, Y., Sakurai, T., Iida, K., Akiyama, K., Satou, M., Toyoda, T., Konagaya, A., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
 Direct Submission
 Submitted (22-MAR-2005) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: mseki@gsc.riken.jp).
 URL: http://range.gsc.riken.jp/, Tel: 81-45-503-9625, Fax: 81-45-503-9586

COMMENT
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720; Seki et al. (2002) Science 296:141-145).
 This clone is in a modified pBluescript vector.
 Please visit our web site (http://range.gsc.riken.jp/) for further details.

FEATURES
 source
 1..1077
 /location="Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAFL22-34-N18"
 /ecotype="Columbia"
 /note="common name: thale cress"
 <1..1077
 /gene="At3g53130"
 <1..906
 /gene="At3g53130"
 /codon_start=1
 /product="Cytochrom P450 -like protein"
 /protein_id="BAD94136.1"
 /db_xref="GI:62319017"

gene
 CDS
 1..1077
 /translation="NSPVEIAVYTTALKAEALRSTDLLPYWKIDALCKIVPRQVRAEKA
 VTLRETVEDILAKKEIVREGERINDEEYVNDADPSILRFLASREVSQVLRDD
 LLSMLVAGHETGSLVTLVLLSNALRAKQEDVRVLEGRNPAPEDIKELKYIT
 RCINESMELYPHPVLIRRAQVDPILLPGNYKVTQGDIMISVYNIHRSSEVWEKAEF
 LPERPDICGALPNETNDFKIPSPSGGRKCVGQDPALMEIALVALVFLQLRVLP
 DQTSIMTTGATIHITNGLYMKVQR"

ORIGIN

Query Match 10.9%; Score 268.6; DB 15; Length 1077;
 Best Local Similarity 68.2%; Pred. No. 1.8e-53;
 Matches 557; Conservative 0; Mismatches 4; Indels 256; Gaps 3;
 QY 1025 TTTTTCAGATCGATGTCATTGTTGAAGATAGTCCCGAGACAGGTGAAAGCTCAAAGGCT 1084
 |||||
 Db 73 TATTGGAAGATCGATGTCATTGTTGAAGATAGTCCCGAGACAGGTGAAAGCTCAAAGGCT 132
 |||||
 QY 1085 GTAACTTTGTATAAGGGAACCTGTTGAAGACCTTATTGCTAAAGTGAAGAAATTTGCGAA 1144
 |||||
 Db 133 GTAACTTTGTATAAGGGAACCTGTTGAAGACCTTATTGCTAAAGTGAAGAAATTTGCGAA 192
 |||||
 QY 1145 AGAAGGCGGAAAGAAATCAATGATGAGGATGTAATGATGCTGACCCAAAGTATCTCTG 1204
 |||||
 Db 193 AGAAGGCGGAAAGAAATCAATGATGAGGATGTAATGATGCTGACCCAAAGTATCTCTG 252
 |||||
 QY 1205 CGTTTCTTGTCTTGAAGCAGAGAGGTTTAAACCTTTTTCCTTAAAGTTTATAAGCAA 1264
 |||||
 Db 253 CGTTTCTTGTCTTGAAGCAGAGAGGTTTAAACCTTTTTCCTTAAAGTTTATAAGCAA 277
 |||||
 QY 1265 TTTGGCCTTTTCATTATCGCATAATCGAAGCTGATGTTGCATTGTGAGGGTTTTTCAGG 1324
 |||||
 Db 278 TTTGGCCTTTTCATTATCGCATAATCGAAGCTGATGTTGCATTGTGAGGGTTTTTCAGG 283
 |||||
 QY 1325 CAAGTGTGAGTTACGGGATGATCTTCTCAATGCTCGTAGCGGGTCATGAACACCTG 1384
 |||||
 Db 284 CAAGTGTGAGTTACGGGATGATCTTCTCAATGCTCGTAGCGGGTCATGAACACCTG 343
 |||||
 QY 1385 GATCTGCTCTCACTTGGACACTTTATCTCTTAAGTAAAGTACCTTAATGATATCTTCTACT 1444
 |||||
 Db 344 GATCTGCTCTCACTTGGACACTTTATCTCTCTTAAGTAAAGTAAAGTAAAGTAAAGTAA 379
 |||||
 QY 1445 TTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTCTGTTCTCATTTTACCTCTTCAA 1504
 |||||
 Db 380 TTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTCTGTTCTCATTTTACCTCTTCAA 379
 |||||
 QY 1505 TTCTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAGACACAAGAAAGTAGACAGAGT 1564
 |||||
 Db 380 TTCTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAGACACAAGAAAGTAGACAGAGT 425
 |||||
 QY 1565 GTTAGAGGAGAAACCCGGCTTTTCGAGGATATAAAGAGTTGAAGTACATCCTCGTTG 1624
 |||||
 Db 426 GTTAGAGGAGAAACCCGGCTTTTCGAGGATATAAAGAGTTGAAGTACATCCTCGTTG 485
 |||||
 QY 1625 TATAAACGAGTCAATGCGCTCTATCTCTCATCTCTCTGTAAGCAATCAAGCTCATCTCTC 1684
 |||||
 Db 486 TATAAACGAGTCAATGCGCTCTATCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 522
 |||||
 QY 1685 TAATTTATTCATGAATAAAATTTTCGATTGATTTGTTCTCTGTTAGGCTTTGATGAAGAG 1744
 |||||
 Db 523 TAATTTATTCATGAATAAAATTTTCGATTGATTTGTTCTCTGTTAGGCTTTGATGAAGAG 536
 |||||
 QY 1745 AGCTCAAGTCTTCGACATTTCTCTCGGAACCTATTAAGTCAATACCGGACAGACATTAT 1804
 |||||
 Db 537 AGCTCAAGTCTTCGACATTTCTCTCGGAACCTATTAAGTCAATACCGGACAGACATTAT 596
 |||||
 QY 1805 GATTTCAGTCTATAACATCCATCGTTCTTCCGAGGTA 1841
 |||||
 Db 597 GATTTCAGTCTATAACATCCATCGTTCTTCCGAGGTA 633
 |||||

RESULT 7

BT012891
 LOCUS Lycopersicon esculentum clone 113997F, mRNA sequence.
 DEFINITION
 BT012891
 ACCESSION
 BT012891.1 GI:47104306
 VERSION
 FLI CDNA.
 KEYWORDS
 Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1967)
Kirkness, E.F., Wang, W. and Vazeille, A.
Direct Submission
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
Location/Qualifiers
1..1967
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="113997P"
/tissue_type="trichome"
/note="THDI89"

ORIGIN
Query Match 10.2%; Score 252; DB 15; Length 1967;
Best Local Similarity 56.8%; Pred. No. 1.6e-49;
Matches 595; Conservative 0; Mismatches 290; Indels 239; Gaps 3;
3;
QY 27 TTCTCTCTTACTCTCTCTCTTCACTGCGAAACCTAGCGGTCTTTTATCACCAAAACC 86
Db 32 TTCTCTCTCTCTCTCTCTCTTCACTAACCTCACCGGACCACTGCTCCGCCCAAAA 91
QY 87 CAAATTCACNTTCCATCAGATCCTCCATTGAGAAACCAACCAAACTCGAGACCA- 145
Db 92 ACCCACTCCAAATCGTTCACAACTCACCAATTAATCTCTTTCGACCAAGAAACCAAC 151
QY 146 ATTTCATCGAAATCCCAATCATGGGTGAGTCCGATGGCTCAACACACTCGTACCC 205
Db 152 CTTCAACTAAGCTACTTCTATGGGTGAGTCCGATGGCTTACTAACTTACCAGTCAAC 211
QY 206 TTCTCTCAGGAAAAACGACGAGTCAGGTATACCAATCGGAAACCGGAGCTCGACGATG 265
Db 212 TTACTTTAGGCCAAAATGATGATTTCTAACATACCCATTCGAGTGTCTGAGCTTGTATG 271
QY 266 TCGCTGATCTCTCGGAGGTCTCTCTTCTACCTCTCTACAAAGGATGATGATGACG 325
Db 272 TTTCGGAATCTCGGCGGTGCTCTTTTCCTTCATGTTATAGATGATGAATTTGTATG 331
QY 326 GACCAATTTACCGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 385
Db 332 GACCTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
QY 386 TAGCTAAACATGTTTGAAGAAATATCCAAAGTACGCTAAAGCTTAGTCTGCTGAGTCT 445
Db 392 TTGCTAAGCATGTTTGAAGAAATATGGAAGATGATGGGAAAGGCTTGTGCTGAGTTT 451
QY 446 CTGAATTTCTATTTGGTTCGGGTTTCGCTATCGCTGAGGAGCTCTTTGACAGTAATTT 505
Db 452 CTGAGTTTGTGTTGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 503
QY 506 CATCTCCTCTATCTCAATTTTGAAGTTTTCGAAATTTGGAAGTAAATGTGACTGTCT 565
Db 504 ----- 503
QY 566 TGTATGATAAGTAACCTAATTTTAGGGTTTAGATTTCCAAATCTTCTATTTGGGCTTACG 625
Db 504 ----- 503
QY 626 TGAAGTCTGATTTTACATAGCGGAGCGTAGGCGGTGTTCCATCGCTTTCACAGGAG 685
Db 504 -----GGCAAGCGGAAGGGCTGTGTTTCCATCTCTTCACAAAGAA 542
QY 686 GTATTTGTCTGATTTGTGAGAGATATTCTGCAAAATGTGACAGAGGCTTGTTCAGAA 745
Db 543 GTACTTGTCAATATAGTTGATCGGCTCTTTTGCAGATGCTGCTGAGAGATGTTGGA 602
QY 746 GTTCAGCCTTATGCGAGAGACGGAAGTGTGGAATATGGAAGCGAAGTCTCTCAGAT 805
Db 603 ACTTTTACCTGATGCAATTTCTGGCTCTGAGTGAATATGGAGGCAAGTTTCTCAACT 662
QY 806 GACACTTGATGTCATTTGGGTTGTCTCTTTTAACTACAAATTTTCGATTTCTTTGACTACTGA 865

Db 663 AACACTTGATTTATGCGCTTGCACCTCTTCAATTAACAATTTTGTATTCCTTACTACTGA 722
QY 866 TAGTCTGTCTATCAAGCTGTTTACACTGCTCTTTAAAGAGCTGAGCTTCTGTTACTCTGA 925
Db 723 CAGTCCAGTTATTTGATGACGTTTACACTGCACTGCACTAAAGAGACGAACTCCGTTCAACTGA 782
QY 926 TCTTCTGCCATATTGGAAGCAAGTTTCTCTGTGTTTTTCTGTGTTGTTGTTGTTGTTGTTG 985
Db 783 TTTGTTGCCATATTGG----- 798
QY 986 GAACAATTTGGAATCTTGTTAATTCAGAGGGTTTGTGTTGTTTTTTCAGATCGATGCAATTG 1045
Db 799 -----CAGATCAAAAGCTTTA 813
QY 1046 TGTAAAGTAGTCCCGAGACAGTGAAGGCTGAAAGGCTGTAACTTTTCATTAAGGGAACCT 1105
Db 814 TGTAAAGTTTCATCCCAAGCAAAATAAGCTGAGATGCAAGTGTCTATTAATCAGACAAACA 873
QY 1106 GTTGAAGACCTTTATTTGCTTAAGTGTAAAGAAATTTGTGAAAGAGAGGCGAAAGAAATCAAT 1165
Db 874 GTTGAAGAACTTTATTTGCAAGTGTGCAAGAGATTGTAGAAACTGAGGGGTGAGAGGATTAAT 933
QY 1166 GATGAGGAGTATGTAATGATGCTGACCCAAAGTATCTGCTGCTTCTGCTTCTGCTGCAAGCAGA 1225
Db 934 GAAGATGAGTACGTGAATGATAGATCCAAAGCATCTTTCGATTTCTGCTTCTGCTAGCCGT 993
QY 1226 GAAGAGGTTTAAACTTTTTCCTT 1249
Db 994 GAGGAGGTTTCAAGTGTACAACCTT 1017

RESULT 8
AP008216_208
WPCOMMENT
Sequence split into 227 fragments LOCUS AP008216 Accession AP008216
Fragment Name Begin End
AP008216_000 1 110000
AP008216_001 100001 210000
AP008216_002 200001 310000
AP008216_003 300001 410000
AP008216_004 400001 510000
AP008216_005 500001 610000
AP008216_006 600001 710000
AP008216_007 700001 810000
AP008216_008 800001 910000
AP008216_009 900001 1010000
AP008216_010 1000001 1110000
AP008216_011 1100001 1210000
AP008216_012 1200001 1310000
AP008216_013 1300001 1410000
AP008216_014 1400001 1510000
AP008216_015 1500001 1610000
AP008216_016 1600001 1710000
AP008216_017 1700001 1810000
AP008216_018 1800001 1910000
AP008216_019 1900001 2010000
AP008216_020 2000001 2110000
AP008216_021 2100001 2210000
AP008216_022 2200001 2310000
AP008216_023 2300001 2410000
AP008216_024 2400001 2510000
AP008216_025 2500001 2610000
AP008216_026 2600001 2710000
AP008216_027 2700001 2810000
AP008216_028 2800001 2910000
AP008216_029 2900001 3010000
AP008216_030 3000001 3110000
AP008216_031 3100001 3210000
AP008216_032 3200001 3310000
AP008216_033 3300001 3410000
AP008216_034 3400001 3510000
AP008216_035 3500001 3610000

AP008216_036 3600001 3710000
AP008216_037 3700001 3810000
AP008216_038 3800001 3910000
AP008216_039 3900001 4010000
AP008216_040 4000001 4110000
AP008216_041 4100001 4210000
AP008216_042 4200001 4310000
AP008216_043 4300001 4410000
AP008216_044 4400001 4510000
AP008216_045 4500001 4610000
AP008216_046 4600001 4710000
AP008216_047 4700001 4810000
AP008216_048 4800001 4910000
AP008216_049 4900001 5010000
AP008216_050 5000001 5110000
AP008216_051 5100001 5210000
AP008216_052 5200001 5310000
AP008216_053 5300001 5410000
AP008216_054 5400001 5510000
AP008216_055 5500001 5610000
AP008216_056 5600001 5710000
AP008216_057 5700001 5810000
AP008216_058 5800001 5910000
AP008216_059 5900001 6010000
AP008216_060 6000001 6110000
AP008216_061 6100001 6210000
AP008216_062 6200001 6310000
AP008216_063 6300001 6410000
AP008216_064 6400001 6510000
AP008216_065 6500001 6610000
AP008216_066 6600001 6710000
AP008216_067 6700001 6810000
AP008216_068 6800001 6910000
AP008216_069 6900001 7010000
AP008216_070 7000001 7110000
AP008216_071 7100001 7210000
AP008216_072 7200001 7310000
AP008216_073 7300001 7410000
AP008216_074 7400001 7510000
AP008216_075 7500001 7610000
AP008216_076 7600001 7710000
AP008216_077 7700001 7810000
AP008216_078 7800001 7910000
AP008216_079 7900001 8010000
AP008216_080 8000001 8110000
AP008216_081 8100001 8210000
AP008216_082 8200001 8310000
AP008216_083 8300001 8410000
AP008216_084 8400001 8510000
AP008216_085 8500001 8610000
AP008216_086 8600001 8710000
AP008216_087 8700001 8810000
AP008216_088 8800001 8910000
AP008216_089 8900001 9010000
AP008216_090 9000001 9110000
AP008216_091 9100001 9210000
AP008216_092 9200001 9310000
AP008216_093 9300001 9410000
AP008216_094 9400001 9510000
AP008216_095 9500001 9610000
AP008216_096 9600001 9710000
AP008216_097 9700001 9810000
AP008216_098 9800001 9910000
AP008216_099 9900001 10010000
AP008216_100 10000001 10110000
AP008216_101 10100001 10210000
AP008216_102 10200001 10310000
AP008216_103 10300001 10410000
AP008216_104 10400001 10510000
AP008216_105 10500001 10610000
AP008216_106 10600001 10710000
AP008216_107 10700001 10810000
AP008216_108 10800001 10910000

AP008216_109 10900001 11010000
AP008216_110 11000001 11110000
AP008216_111 11100001 11210000
AP008216_112 11200001 11310000
AP008216_113 11300001 11410000
AP008216_114 11400001 11510000
AP008216_115 11500001 11610000
AP008216_116 11600001 11710000
AP008216_117 11700001 11810000
AP008216_118 11800001 11910000
AP008216_119 11900001 12010000
AP008216_120 12000001 12110000
AP008216_121 12100001 12210000
AP008216_122 12200001 12310000
AP008216_123 12300001 12410000
AP008216_124 12400001 12510000
AP008216_125 12500001 12610000
AP008216_126 12600001 12710000
AP008216_127 12700001 12810000
AP008216_128 12800001 12910000
AP008216_129 12900001 13010000
AP008216_130 13000001 13110000
AP008216_131 13100001 13210000
AP008216_132 13200001 13310000
AP008216_133 13300001 13410000
AP008216_134 13400001 13510000
AP008216_135 13500001 13610000
AP008216_136 13600001 13710000
AP008216_137 13700001 13810000
AP008216_138 13800001 13910000
AP008216_139 13900001 14010000
AP008216_140 14000001 14110000
AP008216_141 14100001 14210000
AP008216_142 14200001 14310000
AP008216_143 14300001 14410000
AP008216_144 14400001 14510000
AP008216_145 14500001 14610000
AP008216_146 14600001 14710000
AP008216_147 14700001 14810000
AP008216_148 14800001 14910000
AP008216_149 14900001 15010000
AP008216_150 15000001 15110000
AP008216_151 15100001 15210000
AP008216_152 15200001 15310000
AP008216_153 15300001 15410000
AP008216_154 15400001 15510000
AP008216_155 15500001 15610000
AP008216_156 15600001 15710000
AP008216_157 15700001 15810000
AP008216_158 15800001 15910000
AP008216_159 15900001 16010000
AP008216_160 16000001 16110000
AP008216_161 16100001 16210000
AP008216_162 16200001 16310000
AP008216_163 16300001 16410000
AP008216_164 16400001 16510000
AP008216_165 16500001 16610000
AP008216_166 16600001 16710000
AP008216_167 16700001 16810000
AP008216_168 16800001 16910000
AP008216_169 16900001 17010000
AP008216_170 17000001 17110000
AP008216_171 17100001 17210000
AP008216_172 17200001 17310000
AP008216_173 17300001 17410000
AP008216_174 17400001 17510000
AP008216_175 17500001 17610000
AP008216_176 17600001 17710000
AP008216_177 17700001 17810000
AP008216_178 17800001 17910000
AP008216_179 17900001 18010000
AP008216_180 18000001 18110000
AP008216_181 18100001 18210000

[illegible]


```

***, 31 unordered pieces.
AC155564
VERSION 2 GI:58082423
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 205408)
Chen, A.P., Pertea, G., Zheng, L., Cheung, F., Lee, D., Koo, H.,
Utterback, T.R., Feldblyum, T.V., Rabinowicz, P., Fraser, C.M.,
Schubert, K., SanMiguel, P., Ma, J., Pontaroli, A.C., Rohlfing, T.,
Budiman, A., Bedell, J.A., Lakey, N., Barbazuk, B., Bennetzen, J. and
Quackenbush, J.
Consortium for Maize Genomics - BAC skim sequencing and assembly
Unpublished
2 (bases 1 to 205408)
Chan, A.P., Pertea, G., Zheng, L., Cheung, F., Lee, D., Koo, H.,
Utterback, T.R., Feldblyum, T.V., Rabinowicz, P., Fraser, C.M.,
Schubert, K., SanMiguel, P., Ma, J., Pontaroli, A.C., Rohlfing, T.,
Budiman, A., Bedell, J.A., Lakey, N., Barbazuk, B., Bennetzen, J. and
Quackenbush, J.
Direct Submission
Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
9712 Medical Center Dr, Rockville, MD 20850
3 (bases 1 to 205408)
Chan, A.P., Pertea, G., Zheng, L., Cheung, F., Lee, D., Koo, H.,
Utterback, T.R., Feldblyum, T.V., Rabinowicz, P., Fraser, C.M.,
Schubert, K., SanMiguel, P., Ma, J., Pontaroli, A.C., Rohlfing, T.,
Budiman, A., Bedell, J.A., Lakey, N., Barbazuk, B., Bennetzen, J. and
Quackenbush, J.
Direct Submission
Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
9712 Medical Center Dr, Rockville, MD 20850
On Jan 25, 2005 this sequence version replaced gi:57863085.
----- Trace submission -----
Center name: TIGR
Seq_lib_id: ZGEEK
----- Project information -----
Web site: http://www.tigr.org/tdb/tgi/maize/
Contact: maize@tigr.org
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1702: contig of 1702 bp in length
* 1703 1802: gap of unknown length
* 1803 4662: contig of 2860 bp in length
* 4663 4762: gap of unknown length
* 4763 11548: contig of 6785 bp in length
* 11548 15559: gap of unknown length
* 15559 18073: contig of 3911 bp in length
* 18073 22716: gap of unknown length
* 22716 22815: contig of 2414 bp in length
* 22815 35447: gap of unknown length
* 35447 59876: contig of 24329 bp in length
* 59876 61762: gap of unknown length
* 61762 64114: contig of 1786 bp in length
* 64114 71636: gap of unknown length
* 71636 74352: contig of 2252 bp in length
* 74352 78945: gap of unknown length
* 78945 93944: gap of unknown length
* 93944 94944: gap of unknown length
* 94944 95944: gap of unknown length
* 95944 96944: gap of unknown length
* 96944 97944: gap of unknown length
* 97944 98944: gap of unknown length
* 98944 99944: gap of unknown length
* 99944 100944: gap of unknown length
* 100944 101944: gap of unknown length
* 101944 102944: gap of unknown length
* 102944 103944: gap of unknown length
* 103944 104944: gap of unknown length
* 104944 105944: gap of unknown length
* 105944 106944: gap of unknown length
* 106944 107944: gap of unknown length
* 107944 108944: gap of unknown length
* 108944 109944: gap of unknown length
* 109944 110944: gap of unknown length
* 110944 111944: gap of unknown length
* 111944 112944: gap of unknown length
* 112944 113944: gap of unknown length
* 113944 114944: gap of unknown length
* 114944 115944: gap of unknown length
* 115944 116944: gap of unknown length
* 116944 117944: gap of unknown length
* 117944 118944: gap of unknown length
* 118944 119944: gap of unknown length
* 119944 120944: gap of unknown length
* 120944 121944: gap of unknown length
* 121944 122944: gap of unknown length
* 122944 123944: gap of unknown length
* 123944 124944: gap of unknown length
* 124944 125944: gap of unknown length
* 125944 126944: gap of unknown length
* 126944 127944: gap of unknown length
* 127944 128944: gap of unknown length
* 128944 129944: gap of unknown length
* 129944 130944: gap of unknown length
* 130944 131944: gap of unknown length
* 131944 132944: gap of unknown length
* 132944 133944: gap of unknown length
* 133944 134944: gap of unknown length
* 134944 135944: gap of unknown length
* 135944 136944: gap of unknown length
* 136944 137944: gap of unknown length
* 137944 138944: gap of unknown length
* 138944 139944: gap of unknown length
* 139944 140944: gap of unknown length
* 140944 141944: gap of unknown length
* 141944 142944: gap of unknown length
* 142944 143944: gap of unknown length
* 143944 144944: gap of unknown length
* 144944 145944: gap of unknown length
* 145944 146944: gap of unknown length
* 146944 147944: gap of unknown length
* 147944 148944: gap of unknown length
* 148944 149944: gap of unknown length
* 149944 150944: gap of unknown length
* 150944 151944: gap of unknown length
* 151944 152944: gap of unknown length
* 152944 153944: gap of unknown length
* 153944 154944: gap of unknown length
* 154944 155944: gap of unknown length
* 155944 156944: gap of unknown length
* 156944 157944: gap of unknown length
* 157944 158944: gap of unknown length
* 158944 159944: gap of unknown length
* 159944 160944: gap of unknown length
* 160944 161944: gap of unknown length
* 161944 162944: gap of unknown length
* 162944 163944: gap of unknown length
* 163944 164944: gap of unknown length
* 164944 165944: gap of unknown length
* 165944 166944: gap of unknown length
* 166944 167944: gap of unknown length
* 167944 168944: gap of unknown length
* 168944 169944: gap of unknown length
* 169944 170944: gap of unknown length
* 170944 171944: gap of unknown length
* 171944 172944: gap of unknown length
* 172944 173944: gap of unknown length
* 173944 174944: gap of unknown length
* 174944 175944: gap of unknown length
* 175944 176944: gap of unknown length
* 176944 177944: gap of unknown length
* 177944 178944: gap of unknown length
* 178944 179944: gap of unknown length
* 179944 180944: gap of unknown length
* 180944 181944: gap of unknown length
* 181944 182944: gap of unknown length
* 182944 183944: gap of unknown length
* 183944 184944: gap of unknown length
* 184944 185944: gap of unknown length
* 185944 186944: gap of unknown length
* 186944 187944: gap of unknown length
* 187944 188944: gap of unknown length
* 188944 189944: gap of unknown length
* 189944 190944: gap of unknown length
* 190944 191944: gap of unknown length
* 191944 192944: gap of unknown length
* 192944 193944: gap of unknown length
* 193944 194944: gap of unknown length
* 194944 195944: gap of unknown length
* 195944 196944: gap of unknown length
* 196944 197944: gap of unknown length
* 197944 198944: gap of unknown length
* 198944 199944: gap of unknown length
* 199944 200944: gap of unknown length
* 200944 201944: gap of unknown length
* 201944 202944: gap of unknown length
* 202944 203944: gap of unknown length
* 203944 204944: gap of unknown length
* 204944 205408: contig of 1029 bp in length.
Location/Qualifiers
1..205408
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBc0171N08"
1703..1802
/estimated_length=unknown
4663..4762
/estimated_length=unknown
11548..11647
/estimated_length=unknown
15559..15658
/estimated_length=unknown
18073..18172
/estimated_length=unknown
22716..22815
/estimated_length=unknown
35447..35546
/estimated_length=unknown
59876..59975
/estimated_length=unknown
61762..61861
/estimated_length=unknown
64114..64213
/estimated_length=unknown
68253..68352
/estimated_length=unknown
71636..71735
/estimated_length=unknown
74352..74451
/estimated_length=unknown
78945..79044
/estimated_length=unknown

```

/estimated_length=unknown
82989. .83088
/estimated_length=unknown
84258. .84357
/estimated_length=unknown
88369. .88468
/estimated_length=unknown
103421. .103520
/estimated_length=unknown
111901. .112000
/estimated_length=unknown
139515. .139614
/estimated_length=unknown
141413. .141512
/estimated_length=unknown
146354. .146453
/estimated_length=unknown
150399. .150498
/estimated_length=unknown
152461. .152560
/estimated_length=unknown
153780. .153879
/estimated_length=unknown
167813. .167912
/estimated_length=unknown
170603. .170702
/estimated_length=unknown
198609. .198708
/estimated_length=unknown
201018. .201117
/estimated_length=unknown
204280. .204379
/estimated_length=unknown

ORIGIN

Query Match 6.0%; Score 146.8; DB 14; Length 205408;
Best Local Similarity 65.0%; Pred. No. 1.9e-24;
Matches 217; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 632 CTGATTTTACATAGCGGAGCGGTAGACGGTGGTTCATCGCTTCACAGAGGTATTT 691
Db 108385 CTGGTATGTAATTAGTGAGACGCAAGTGTGTGCTTCATTCGATATAATTTAT 108444

QY 692 GTCTGTGATGTGGAGAGATATCTGCAATGTGACAGAGCGCTTGTGAGAAGTTGCA 751
Db 108445 GTCCATTATCGTGGAAAGGTATTTGTAAATATGATGAGAGATTAATAGACAACTTGA 108504

QY 752 GCCTTATGACAGACGGAAGTGTGTAATATGGAAGCGAAGTTCCTCAGATGACACT 811
Db 108505 GCATATGCTTTGAGTGGGGAACCTGTCAATATGAAGCAAGTTTCTCAGTTGACAT 108564

QY 812 TGATGTCAATGGTGTCTCTTTTAACTACAATTTGCAATTCCTTGAAGTCTGATAGTCC 871
Db 108565 GGATGTATGTGTTGTCTCATTTGTTCACTACACTTTTATTCCTCAGATGACAGTTC 108624

QY 872 TGTCAATGAAGCTGTTTACATGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTCT 931
Db 108625 TAGCATTTGTTGTTTATATGCTCACTCAAGAGCAGAGATTCGTTCTACAGATCTTT 108684

QY 932 GCATATTGGAAGCAAGTTTCTCTGTTTCTTTC 965
Db 108685 TGCCATACTAGAAGTGTCTCTCTTTAACTTC 108718

RESULT 13

LOCUS AY142017 1788 bp mRNA linear PLN 02-SEP-2002
DEFINITION Arabidopsis thaliana Atlg31800/68069_m00159 mRNA, complete cds.
ACCESSION AY142017
VERSION AY142017.1 GI:22655377
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1788)
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Heuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Arabidopsis ORF clones

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 1788)
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hayashizaki, Y., Heuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A. and Ecker, J.R.

TITLE Direct Submission

JOURNAL

Submitted (15-AUG-2002) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Heuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers
1. .1788
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U14026"
/ecotype="Columbia"
/note="This clone is in pUNI 51"

CDS

1. .1788
/note="putative cytochrome P450"
/codon_start=1
/product="Atlg31800/68069_m00159"
/protein_id="AAM98281.1"
/db_xref="GI:22655378"

ORIGIN

Query Match 3.7%; Score 91; DB 15; Length 1788;
Best Local Similarity 60.9%; Pred. No. 9e-11;
Matches 148; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 701 TGTGGAGAGATATTCTGCAATGTGCAGAGAGGCTTGTGAGAGTTCAGCCTTATGC 760
DB 645 TATGATTATTTATTCGGAGAGCTTCAGATAGGCTTTGTCAAGCTTGATGCTGCTGC 704

QY 761 AGAAGACGGAAGTGTGTAATATGGAAGCGAAGTCTCTCAGATGACACTTGTATCAT 820
DB 705 ATTGAAGGCGGAAGATAGAGATGGAATCACTCTCTCTCTGTTGACACTTATATAT 764

QY 821 TGGGTTGTCTCTTTTAACTACAATTCGATCTTTTGACTACTGATGCTGCTATGCA 880
DB 765 TGGCAAGGCGGTTTTCAATTCAGACTTTGACTCCCTTACTATGATACCGGTTGATCGA 824

QY 881 AGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTGTGTTCTACTGATCTTCTGCCATATG 940
DB 825 GCGAGTGTACACTGTTCTTAAGAGAGCTGAGACAGAGTGTTCACCTATTCCTGTTG 884

QY 941 GAA 943
DB 885 GGA 887

RESULT 14
LOCUS AY058173 2017 bp mRNA linear PLN 04-NOV-2001
DEFINITION Arabidopsis thaliana Atlg31800/68069_m00159 mRNA, complete cds.
ACCESSION AY058173.1 GI:16648792
VERSION FLI CDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 2017)
AUTHORS Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1017)
Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Banh, J.,
Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission
Submitted (28-AUG-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGE (SSP) Consortium members carried out the
sequencing and annotation of the RAPL cDNAs: Cheuk, R., Chen, H.,
Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L.,
Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,

Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
source

1..2017
/location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAF109-10-L12 (R14026)"
/ecotype="Columbia"
1..38
39..1826
/note="putative cytochrome P450"
/codon_start=1
/product="Atlg31800/68069_m00159"
/protein_id="RAL25587.1"
/db_xref="GI:16648793"
/translation="MAMAPPLSYTPTITVKPVYSSRRNFVVFSSSSNGRDPLENSV
PNVKSLEKLEERERASIASGATVRSKSPSTVNGLSKIGIPSNVLDPMFD
WTGSDQDPKVPKAGSIQAVRNEAFFIPLYELFTYGGIFLRTFGPKSFLIVSDPSI
AKHILKONAKYSGITAEILDVFMGKGLIPADGEIWRRRRAIYPALHQRKVAAMIS
LFGSADRLCQKDAALKEBEMESLFSRLTLDIIGKAVNFDVDTGVTIEA
VYTVLRADRSVSPIPWDIPWKDISPRKQVATSKLINDTLDLIATCKRWVEE
RELQFHEHYMNERDPSILHFLASGDVSKQLRDLMTLMIAHGETSAVLTTWTFYL
LTTFSSVVKLQEEVDVSGIDRPFQIDMKKLKYTRVWNSRLRLYPQPVILRRSID
NDILGEYPIKKEGDIPIFISVWNLHRSPLHWDDAEKFNPERWPLDGNPNETNQFSLP
FGGPRKICIDMFASFENVAIAMLIRFPNFQIAPGAPPVKMTTGATHTTBLGLTV
TKTKPLDIPSPILPMDTSRDEVSSALS"
1827..2017

3'UTR
ORIGIN

Query Match 3.7%; Score 91; DB 15; Length 2017;
Best Local Similarity 60.9%; Pred. No. 8.8e-11;
Matches 148; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 701 TGTGGAGAGATATTCTGCAATGTGCAGAGAGGCTTGTGAGAGTTCAGCCTTATGC 760
DB 683 TATGATTATTTATTCGGAGAGCTTCAGATAGGCTTTGTCAAGCTTGATGCTGCTGC 742

QY 761 AGAAGACGGAAGTGTGTAATATGGAAGCGAAGTCTCTCAGATGACACTTGTATCAT 820
DB 743 ATTGAAGGCGGAAGATAGAGATGGAATCACTCTCTCTGTTGACACTTATATAT 802

QY 821 TGGGTTGTCTCTTTTAACTACAATTCGATCTTTTGACTACTGATGCTGCTATGCA 880
DB 803 TGGCAAGGCGGTTTTCAATTCAGACTTTGACTCCCTTACTATGATACCGGTTGATCGA 862

QY 881 AGCTGTTTACACTGCTCTTAAAGAGCTGAGCTTGTCTACTGATCTTCTGCCATATG 940
DB 863 GCGAGTGTACACTGTTCTTAAGAGAGCTGGAAGACAGAGTGTTCACCTATTCCTGTTG 922

QY 941 GAA 943
DB 923 GGA 925

RESULT 15
LOCUS AY056446

DEFINITION Arabidopsis thaliana Atlg31800/68069_m00159 mRNA, complete cds.
ACCESSION AY056446
VERSION AY056446.1 GI:15912336
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 2057)

Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, H.,
Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.

Arabidopsis cdna clones

Unpublished

2 (bases 1 to 2057)

Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, H.,
Bowser, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y.,
Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M.,
Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K.,
Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C.,
Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K.,
Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (10-SEP-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Chen, H., Chen, H.,
Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Ban, H., Bowser, L.,
Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,
Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A.,
Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
Davis, R.W., Theologis, A., and Ecker, J.R.

Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

Location/Qualifiers

1. .2057

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosomes="1"

/clone="RAFL09-33-D23 (R14331)"

/ecotype="Columbia"

1. .33

34. .1821

/notes="putative cytochrome P450"

/codon_start=1

/product="Atlg31800/68069_m00159"

/protein_id="AAL08302.1"

/db_xref="GI:15912337"

/translation="MAMAPLSVTPITVKPVTSRSPNVFSSSSNGRDPLENSV
PWGSKLEKQEKRAELSAIASGAFVTRKSSFPSTYKGLSKIGIPSNVLDWFD
PTGSDQDPKVPKAGSIQAVNEAFPIFLYELFYGGIFRLTFGPKSFLVSDPSI
AKHILKDNKAYSLGAILDFVMGKGLIPADGEIWRRRRAIVPALHOKYVAAMIS
LFGEASDRCLKDAALKEEVESELSRLTLDIIGKAVNFYDFSLTNDTGVIEA
VTVLREADRSVPVWDIPWKDIPSPORKVATSLKINDLTLATCKRMVEE
BELQFHEEYNERDPSILHLLASGDDVSKOLRDLMTMLIAGHETSAVLTWTFYL
LTTPSVVAKLQEVDSVIGDRPTIQDMKKLKYTRVNESLRLYPQPPVLRSLD
NDILGYPIKRGEDIFISVNLHRSPLHWDIAEFNRPWPLDGPFPNETNQNFSLP
FGGPRKCIQDMFASFNVAIAMLIRRFNFQIAPGAPPVKKMTTGATITHTGLKLT
TKRTKPLDIPSPVLPMDTSRDEVSSALS"

1822. .2057

3'UTR

misc_difference 1989

/notes="compared with genome"

ORIGIN	/replace=" "
Query Match	3.7%; Score 91; DB 15; Length 2057;
Best Local Similarity	60.9%; Pred. No. 8.8e-11;
Matches 148; Conservative	0; Mismatches 95; Indels 0; Gaps 0;
QY	701 TGTGAGAGAGCTATTCTGCAATGTGCAGAGAGGCTTGTTCAGAAAGTTGCAGCCCTTATGC 760
Db	760 TATGATTAGTTTATTCGGAGAAAGCTTCAGATAGGCTTTGTTCAGAAAGCTTGTATGCTGCTGC 737
QY	761 AGAAGACGGAAGTGTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACATTGATGTCAT 820
Db	738 ATTGAAGGGGGAAGAGTAGAGATGGAATCACTCTTCTCTCGTTTGACACTTGAATATTAT 797
QY	821 TGGGTGTCTCTTTTAACTACAATTTCGATTCCTTTGACTACTGATAGTCCCTGTCTATTGA 880
Db	798 TGGCAAGCGGTTTTCAATTACGACTTTTGACTCCCTTACTAATGATACCGGTGTGATCGA 857
QY	881 AGCTGTTTACACTGCTCTTTAAAGAGCTGAGCTTCGTTCTACTGATCTTTCTGCCCATATTG 940
Db	858 GCGAGTGACACTGTTCTTAAGAGAGCTGAAGACAGAAAGTGTTCACCTATTTCCTGTTTG 917
QY	941 GAA 943
Db	918 CGA 920
Search completed: January 13, 2006, 04:30:26	
Job time : 12354 secs	

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 13, 2006, 00:04:14 ; Search time 434 Seconds
(without alignments)
10104.250 Million cell updates/sec

Title: US-10-751-235-5
Perfect score: 2467
Sequence: 1 atggagcttcactcttttc.....tgaaggtgagccaaaggtaa 2467

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73.2	3.0	1831	3	US-08-948-564-15
2	57.8	2.3	404	3	US-09-615-192A-156
3	57.8	2.3	404	3	US-09-169-789-156
4	49.8	2.0	1141	3	US-09-806-708B-22
5	49.2	2.0	1141	3	US-09-806-708B-22
6	48.8	2.0	832	3	US-09-621-976-2813
7	46.8	1.9	7218	2	US-08-232-463-14
8	44.6	1.8	1664976	3	US-08-916-421B-1
9	44.6	1.8	1664976	3	US-09-692-570-1
10	41.8	1.7	43192	3	US-09-949-016-15466
11	41.4	1.7	601	3	US-09-949-016-130943
12	41.4	1.7	67899	3	US-09-949-016-15432
13	41.4	1.7	67902	3	US-09-949-016-15432
14	40	1.6	832	3	US-09-621-976-2813
15	39.6	1.6	601	3	US-09-949-016-88266
16	39.6	1.6	601	3	US-09-949-016-88266
17	39.6	1.6	49011	3	US-09-949-016-14221
18	39.6	1.6	168394	3	US-09-949-016-13002
19	39.4	1.6	1125	3	US-09-248-796A-6024
20	39.4	1.6	1308	3	US-09-902-540-3557
21	39.4	1.6	19019	3	US-09-902-540-1171
22	39	1.6	672	3	US-09-248-796A-1435
23	38.8	1.6	58407	3	US-08-916-421B-2
24	38.8	1.6	58407	3	US-09-692-570-2

c	25	38.2	1.5	7218	2	US-08-232-463-14	Sequence 14, Appl
	26	38.2	1.5	135476	3	US-09-949-016-12611	Sequence 12611, A
	27	38.2	1.5	135476	3	US-09-949-016-14413	Sequence 14413, A
c	28	38	1.5	601	3	US-09-949-016-200471	Sequence 200471, A
c	29	38	1.5	32392	3	US-09-662-254B-27	Sequence 27, Appl
	30	37.2	1.5	113966	3	US-09-949-016-12277	Sequence 12277, A
	31	37.2	1.5	113967	3	US-09-949-016-17051	Sequence 17051, A
	32	37	1.5	268449	3	US-09-949-016-17244	Sequence 17244, A
	33	36.8	1.5	2078	3	US-09-949-016-2102	Sequence 2102, Ap
c	34	36.8	1.5	3150	3	US-10-018-730A-3	Sequence 3, Appl
c	35	36.8	1.5	4312	3	US-10-104-047-468	Sequence 468, App
c	36	36.8	1.5	74177	3	US-09-949-016-11988	Sequence 11988, A
c	37	36.8	1.5	74177	3	US-09-949-016-17388	Sequence 17388, A
	38	36.6	1.5	1584	3	US-09-351-223-1	Sequence 1, Appl
	39	36.4	1.5	601	3	US-09-949-016-92564	Sequence 92564, A
	40	36.4	1.5	601	3	US-09-949-016-92565	Sequence 92565, A
	41	36.4	1.5	1473	3	US-08-277-031B-11	Sequence 11, Appl
	42	36.4	1.5	1473	3	US-08-277-031B-12	Sequence 12, Appl
	43	36.4	1.5	1473	3	US-08-277-031B-13	Sequence 13, Appl
	44	36.4	1.5	1829	2	US-08-201-118-8	Sequence 8, Appl
	45	36.4	1.5	1829	2	US-08-238-821B-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-948-564-15
; Sequence 15, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminszky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948.564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20..1747
US-08-948-564-15

Query Match 3.0%; Score 73.2; DB 3; Length 1831;

	Best Local Similarity	59.98;	Pred No. 1.3e-10;	Mismatches	Conservative	0;	Indels	0;	Gaps	0;
QY	730	AGAGGCTTTGTGGAGAAGCTTGACGCCCTTATCGAAGAAGACGGAGTCGTGTGAATATGGAA	789							
Dδ	612	ATAAGCTTCTTGAAGGAGAGGGTTATGATGGACCTGCATCAATTGAAATTTGGATCTTGAGG	671							
QY	790	CGAAGTTCCTCTCAGATCACACTTTCATGCTCATTTGGGTTGTCTCTTTTTAACTPACAATTTCC	849							
Dδ	672	CAGAGTTTTCTAGTTTGGCTCTTGATATTATTGGGCTTGGTGTCTCAACTATGACTTTG	731							
QY	850	ATTCTTTGACTACTGATAGTCTGTGCTATGAGCTGTTTACACTGCTCTCTTAAAGAACCTG	909							
Dδ	732	GTTCCTGTGCAACAAGAATCTCCAGTTATTTAAGGCAGCTCTATGGCACICTCTTTTGAAGCTG	791							
QY	910	AGCTTCGTTCTACTGATCTTTCTGCCAATATGGAA	943							
Dδ	792	AACACAGATCCACTTTCTACATTTCCATATTTGGAA	825							

RESULT 2
US-09-615-192A-156
; Sequence 156, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkaia, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C4U

	Query Match	2.3%	Score 57.8;	DB 3;	Length 404;
	Best Local Similarity	58.4%;	Pred. No. 1.8e-06;		
	Matches	101;	Conservative	0;	Mismatches 72; Indels 0; Gaps 0
Qy	771	AGTCTGTGATATGGAGCGAAGTTCTCTCAGATGACACITGTAGTCTATTGGGTTGCT	830		
Db	25	ATTGAGGTTGACATGGAGGCGAGAATTTTCAAATCTAGCTTTGGACATTATTGGATTGTT	84		
Qy	831	CTTTTTTAACACAAATTTTCGATTCTTTTGACTACTGTAGTCTCTGTCTATTGAAGCTCTTTAC	890		
Db	85	GTATTTTAACATGATTTTGGATCCGTTACTCGAAGATCACCAGTAATCAAGCAGCTCTAT	144		
Qy	891	ACTGCTCTTAAAGAAGCTGAGCTTCGTTCTACTGTACTCTTGTGCCATATTTGGAA	943		
Db	145	GGTACATTGTTTGAAGCTGAGCATAGATCAACTTTTATACATACCACTACTGGAA	197		

RESULT 3
US-09-169-789-156
; Sequence 156, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C2

```

; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: US 08/713,000
; EARLIER FILING DATE: 1996-09-11
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 156
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-156

```

	Query Match	2.3%	Score 57.8	DB 3	Length 404	
	Best Local Similarity	58.4%	Prod. No. 1.8e-06			
	Matches 101	Conservative	0	Mismatches 72	Indels	Gaps 0
Qy	771	AGTGCTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGATGTCTCATTCGGGTGTGCT	830			
Db	25	ATTGAGGTTCATGTCGAGCGAGAAATTTTCAAACTCTAGCTTTGGACATTTGGATTGTGT	84			
Qy	831	CTTTTTAACTACAAATTTTCGATTCTTTTGACTACTGTAGTCTCTGTCTATTGGAAGCTGTTTAC	890			
Db	85	GTAITTAATATGATTTTGGATCCGTTACTCGAAGATCACCAGTAATCAAGGCAGTCTAT	144			
Qy	891	ACTGCTCTTAAAGAGCTGAGGCTTCGTTCTACTGTACTTCTGCCATATTGGAA	943			
Db	145	GGTACATTTGTTGAAGCTGTGAGCATAGATCAACCTTTTATCATACCATTCTGGAA	197			

RESULT 4
 US-09-806-708B-22/c
 ; Sequence 22, Application US/09806708B
 ; Patent No. 6784342
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
 ; FILE REFERENCE: 4810-58741
 ; CURRENT APPLICATION NUMBER: US/09/806,708B
 ; CURRENT FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: US 60/147,133
 ; PRIOR FILING DATE: 1999-08-04
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: Patent In version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 1141
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; NAME/KEY: promoter
 ; LOCATION: (1)..(1141)
 ; OTHER INFORMATION: Consensus sequence of A.t., L.a., and B.n. PAEL promoters
 ; US-09-806-708B-22

Query Match	2.0%;	Score 49.8;	DB 3;	Length 1141;
Best Local Similarity	10.7%;	Pred. No. 0.0007;		
Matches 92;	Conservative 319;	Mismatches 439;	Indels 9;	Gaps 2;
QY	1047	GTAAGATAGTCCCGAGACAGCGTGAAGAGCTGAAAGGCTGTAACTTTGTATGAAGGAAACTG	1106	
DB	954	GBWADTAGKCNKNNNNNTVDVRRMAMKAKNNNNNNAYWTAACYNRAAATNNKMAATHMKW	895	
QY	1107	TGGAAGACCTATTCTAGCTTAAAGAAATGTTCGAAAGAGAGGCGAAAGAAATCAATG	1166	
DB	894	TGAGASKRRTRHRTKCRTKNNNNNNARTVYWHHAARRVWNAWWTNTNNNNNNNNN	835	
QY	1167	ATGAGGAGTATGTAATGATGTCGCCCAAGTATCTCGGTTTCTGCTTGCAGACGAG--	1224	
DB	834	ACRNRTRTWABWKHSCVNNNNNNNNNNNNNTWCHYTTANABBCYRANNNNNAARMATC	775	
QY	1225	AGAAGAGGTTTAAACTTTTTCCTTAAGTTTATAAGCAAAATTTGGCTTTTCATTATCGCA	1284	

Db 774 NNYMHAAVTTTHDWCYKTMNTYWDMMTMBTTTTTRNTTSTNNNNNNNNMMACTNN 715
 QY 1285 TAATCGAAGCTGATGTGTCAGTGTTCAGGAGTTTTCAGGTATCAAGTGTGCGAGTACCGGAT 1344
 Db 714 NNNNNMKAYYAHATNMGCGWNNNTDARRNTNTVMRRNMTNTKRWYSTTRRHHTGAT 655
 QY 1345 GATCTTCTCAATGCTGCTAGCGGTGATGAACCACTGGATCTGCTCACTTGGACA 1404
 Db 654 NNNNNNNNNNNNNNSCCTTRMTMTMTKMGDGMTVKRKKWRDITCTCTYVDVWADSVV 595
 QY 1405 CTTTATCTCTCAAGTAAAGTCACTTAATATCTCTTCTACTTGTCTATGCTAGAGATTTA 1464
 Db 594 WYVANMRCEDVTYNTNTYCKSYAHSYWNSNMMWYRYRARNWSSWARTTNNWMM 535
 QY 1465 CTTGGAAGGAGCTTCTGTTCTCATTTAATCACTCTTCAAAATCTCTATGTTCAAGAACT 1524
 Db 534 SGBVRMRWAGTMMWRHNNNNNTDTRYMMWKRARBTTTVYDSMCNAKSMRGNWRAMK 475
 QY 1525 CATCT-----GCATTAAGGAAGCACAAAGGAAGTAGACAGAGTGTAGAGGAAGA 1577
 Db 474 MWAANNDAGADHWTYMMGNNTMMRRAMWMMMAWCCRAYCCNNNNNRACVHHKHM 415
 QY 1578 AACCGCGCTTTCAGGATATAAGGAGTTGAAGTACATCACTCGTTGTATAAAGAGTCA 1637
 Db 414 RWTYKYMKAACANNNBKAMYRVAMMYSRDITNTDMMWTSDHBWHYTYDYTMRAW 355
 QY 1638 ATGCGTCTCTATCTCTCATCTCTGTGAAGCAATCAAGCTCATCTCTCTAATTTATTCATGA 1697
 Db 354 NNNNNNNRBCKTTSMWMMHMTHTCTYGNNTWGSAYBMAAMSWAAGASBNVTYNWCM 295
 QY 1698 ACTAAATTTCTGATGATTTGTTCTGCTGAGTCTTGTATGAAGAAGAGCTCAAGTCTCT 1757
 Db 294 RMTYMGKTNNNNNNKAWYRTKTVACNNRYYDTAVMTBKRYKYCIAYBMYBYM 235
 QY 1758 GACATCTTCTCGGGAACATATAAGTCAATACCGGACAGACATATGATTTCACTCTAT 1817
 Db 234 GCHHWBRBRHRSNNMMWVKRYMVSWHYHMYRBYKABAVGCCNNNWKDRMAHHW 175
 QY 1818 AACATCCATCGTCTTCCGAGTACAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1877
 Db 174 CATNNNNMMWYAYMHHMKKGAATNNKTAABRDHBAHVKTYYWRYDYWCAMCWNNA 115
 QY 1878 GGGGAGCCTAATCTCTCTC 1896
 Db 114 KAKVETAMKMMYYTDRYV 96

RESULT 5

US-09-806-708B-22
 ; Sequence 22, Application US/09806708B
 ; Patent No. 6784342
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of British Columbia
 ; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
 ; FILE REFERENCE: 4810-58741
 ; CURRENT APPLICATION NUMBER: US/09/806,708B
 ; CURRENT FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: US 60/147,133
 ; PRIOR FILING DATE: 1999-08-04
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: Patent version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 1141
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; NAME/KEY: promoter
 ; LOCATION: (1)..(1141)
 ; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
 US-09-806-708B-22

Query Match 2.0%; Score 49.2; DB 3; Length 1141;
 Best Local Similarity 11.3%; Pred. No. 0.001;

RESULT 6
 US-09-621-976-2813

Matches 110; Conservative 354; Mismatches 499; Indels 10; Gaps 3;
 QY 840 TACAAATTCGATCTTTGACTACTAGTCTCTGTCATTTGAAGCTGTTTACACTGCTCTT 899
 Db 60 TWARWYCKYRRYNNKSRWKGWYKKWYBCANNSTSBRYHARRWKMKTAYBMTWTKWG 119
 QY 900 AAGAAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTGGAAGCAAGTTTCTCTGTCT 959
 Db 120 KTGMRHYRWRAMBDDTYDHHYVTAJNNNAWTTTMCMDKDDKRTRWMMKKNNNATGMDDT 179
 QY 960 TTTTCTCTGTTGTTGATTTGTTGGAACAATTCGATTTCTTTGTTAAATTTGAGAGGTTTG 1019
 Db 180 KXHMNNNGCBVTVMVRYKDRDWSBKRMVYGBWKNWSYDVTVYVWVWDMCKRKYR 239
 QY 1020 GTTGTGTTTTTTCAGATCGATGTCATTTGTGTAGATAGTCCCGAGACAGAGTGAAGCTGAAA 1079
 Db 240 RWVTRGRMRNTVMVAWBTAHRRRYNNGWMTBAMAYRRWTTNNNNNNNAKAMCKRKYWGNR 299
 QY 1080 AGGCTGTAACTTTGATAAGGGAACCTGTTGAAGACCTTATTCCTAAGTCTAAGAATTTG 1139
 Db 300 A-BVNSTCTTWKSKTTKVRTSCWANNCRAGDANKDHKWKWSAAMGVVYNNNNNNNTYK 358
 QY 1140 TCGAAGAGAAGGC--GAAGAATCAATGATGAGGAGTATCTAAATGATGCTGAGCCCAA 1196
 Db 359 KARHBRADWVWHSAAKWKHANAHYSRKWTBYRKTWVNNNGTTTMMKRWMAWTKWD 418
 QY 1137 GTATCTCGTGTCTTCTGCTGCAAGCAGAGAGGTTTAAACTTTTTTCTTAAAGTTTAA 1256
 Db 419 MDWBTGTYNNNNNGRTTYTGWTKNKKOMWTYKWKANNCKWRAWDHKTCTHNTTMMKMTY 478
 QY 1257 TAAGCAAAATTTGGCCCTTTTCATATTCGCAATATCGAAGCTGATGTTGCTATGTGAGGTTT 1316
 Db 479 WNNCYKSMNTGKSHRBAAYVYTWVWRRYAHANNNNNDYWWKACTWYKYBVCWKWN 538
 QY 1317 TCAGTATCAAGTGTGCGAGTTACGGGATGATCTTCTCTCAATGCTGCTAGCGGTCATCA 1376
 Db 539 NYAAWYTKSSWNTSYRYRWKTNNSWRSD-----TRSMGRANNYARABHYGYKWNTR 592
 QY 1377 AACCACTGATCTGTCTCTCACCTGAGACACTTTATCTCTAAGTAAGTACCTTAATGAT 1436
 Db 593 WMBKSHTWBHRAGAAYHWMEMBYBAKCHMKAWYKAKYAGAGGSGNNNNNNNNNNNN 652
 QY 1437 CTTCTACTTCTGATGATGAGAAATTTACTTGGATGGGAGCTTCTCTGTTCTCATTTACC 1496
 Db 653 NNATCARDYYAASRWYAMANAAMWYKBAANNAYYTHANNWMCWNNATDTRRTWKN 712
 QY 1497 TCTTCAATCTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAGCAAGAAGTGA 1556
 Db 713 NNNNAGTWKNNNNNNNAKASAAKNVAAAAYKAAKHWKRWKAWRGHADAABTTDKR 772
 QY 1557 GACAGAGTGTAGAGGAAGAAACCCGCTTTTCGAGGATATAAAGAGGTTGAAGTACATC 1616
 Db 773 NGAYTKYTTTTNNNTYRGVVTNTAARDGWANNNNNNNNNNNNNNNGSDMWVWYWAYNY 832
 QY 1617 ACTCGTTGTAAGAGGAGTCAATGGGTCTCATCTCATCTCTCTGTAAGCAATCAAGCT 1676
 Db 833 GTNNNNNNNNNAYAWMTNKWYTTTDRWRBAYTNNNNNNNNRMAYYGAYADDYAYYMSDTC 892
 QY 1677 CATCTCTTAATTTATCATGAATAAATTTTCTGATGATTTGTTCTCTGTTAGTCTGTTG 1736
 Db 893 DAWMKDATKNNNATTTNNGRTAWRTNNNNNNMMTKTYBYBHAAMNNNNNNNGHCTAHTW 952
 QY 1737 ATAAGAAGAGCTCAAGTTTCTGACATCTTCTCTGGAACTATAAGGTCAATACCGGACAA 1796
 Db 953 VKATKTKGCMWNCCTTTCRYKNNCTWYTWMTTTRTWTAAATRWKTNATGSMTRCNAT 1012
 QY 1797 GACATTATGATTT 1809
 Db 1013 GWKNNYWTGKTT 1025


```
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURES:
  NAME/KEY: misc_feature
  LOCATION: (28222)..(28222)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (28257)..(28258)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (84773)..(84773)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (84808)..(84808)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (84812)..(84812)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (98120)..(98120)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (98159)..(98159)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (98239)..(98239)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (98266)..(98266)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (98343)..(98343)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (103998)..(103998)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (148948)..(148948)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (163385)..(163385)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (191989)..(191989)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (191995)..(191995)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (231980)..(231980)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (234187)..(234187)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (234220)..(234220)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (234814)..(234814)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (309398)..(309398)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (309418)..(309418)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (312837)..(312837)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (312993)..(312993)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (319226)..(319226)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (559167)..(559167)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (559241)..(559241)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (600992)..(600992)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (622708)..(622708)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (657081)..(657081)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (657203)..(657203)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (674435)..(674435)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (682442)..(682442)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (713652)..(713652)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (741684)..(741684)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (779455)..(779455)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (779676)..(779676)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (855539)..(855539)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (871619)..(871619)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1084830)..(1084830)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1096846)..(1096846)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1119881)..(1119881)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1130881)..(1130881)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1310988)..(1310988)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1313224)..(1313224)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1349473)..(1349473)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1349491)..(1349491)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1470091)..(1470091)
  OTHER INFORMATION: n equals a, t, c, or g
  NAME/KEY: misc_feature
  LOCATION: (1569020)..(1569020)
```

```
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
;
US-08-916-421B-1
Query Match 1.8%; Score 44.6; DB 3; Length 1664976;
Best Local Similarity 50.2%; Pred. No. 1.4;
Matches 110; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 1060 GAGACAGGTGAAGCTGAAAGGCTGTAACCTTTGATTAAGGGAACCTGTTGAGACCTTAT 1119
Db 1435967 GAGAGAAATGACATCAAGAGGAATTAATGGAATTAATTGAAGCAATTAATATACGTC 1436026
QY 1120 TGCTAAGTCTAAGAAATTTGCGAAGAGAGAGCGCAAGAAATCAATGATGAGGAGTATGT 1179
Db 1436027 TGAAGAACTGGAAGAGATTTGTCATGTTAAGGACCAAGATCATTTTAAGAGAGTAG 1436086
QY 1180 AATGATGTCACCAAGTATCCTGCTTTCTTCTTCTGCAAGCAGAGAGAGGTTTAAAC 1239
Db 1436087 GATTATTGATGTTCAAGGAGATGAAGGATATATATTAGAAGGAAGAGGATGGA 1436146
QY 1240 TTTTTCCTTAAGTTTATTAAGCAAAATTTGGCCCTTTTCATT 1278
Db 1436147 GATAAAGCAAGATTAATTGTTTAAAAAGGCTATAAATT 1436185
RESULT 9
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; PATENT NO. 6797466
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
```

```
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
;
;
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (319236)..(319236)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
```

```
;
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
;
;
; Query Match 1.8%; Score 44.6; DB 3; Length 1664976;
; Best Local Similarity 50.2%; Pred. No. 1.4;
; Matches 110; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
;
; QY 1060 GAGACAGGTGAAAGCTGAAAAGCGCTGTAACCTTTGATAAGGGAACCTGTTGAGACCTTAT 1119
; Db 1435967 GAGAGAAATGAGCATCAAAAGAGGAAATTAATGGAATTAATTGAAGCAATTAATATATACGTC 1436026
;
; QY 1120 TGCTAAGTGTAAAGAAATTTGCGAAGAGAGAGCGGAAAGCAATCAATGATGAGGATATGT 1179
; Db 1436027 TGAATAAAGCTGAAGAGATTTTCATGTTAAGACCAAGATCATTGTTAAGAGAGTAG 1436086
;
; QY 1180 AAATGATGCTGACCAAGTATCTCTGCTTTCTTGTTCGACGAGAGAGAGGTTTAAAC 1239
; Db 1436087 GATTATTGATGTTCAAGGAGATGAAGGATATATTAGAAGGAGGAGGATCGAAA 1436146
;
; QY 1240 TTTTTCCTTAAGTTTATAAGCAAAATTTGCGCTTTTCATT 1278
; Db 1436147 GATAAAGGCAAGATTAATTGTTAAAAAAGGCTATAAATT 1436185
;
;
; RESULT 10
; US-09-949-016-15466
; Sequence 15466, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15466
; LENGTH: 43192
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15466
;
; Query Match 1.7%; Score 41.8; DB 3; Length 43192;
; Best Local Similarity 59.9%; Pred. No. 1.2;
; Matches 88; Conservative 0; Mismatches 57; Indels 2; Gaps 1;
;
; QY 770 AAGTCCTGTGATATGGAAGCGAAGTCTCTCAGATGACACTTGATGCTGGGTGTC 829
; Db 1863 AAGTCCTGGAGTATAGCGGTGAGCCACGTCCTGGCCCTAGATTCATTGTTATCTC 1922
```

QY 830 TCTTTTAACTACAA--TTTCGATCTTTGACTACTGATAGTCCCTGTCATTTGAAGCTGTT 887
Db 1923 TGTCTTAAGCAAAAATCTTTTCTCTTAGACTTCAGATTTTCTGCTTGATCCAGC 1982
QY 888 TACACTGCTCTTAAAGAAGCTCAGCTT 914
Db 1983 TGGACTGCTCAAGAAGAAATGSCCTT 2009

RESULT 11
US-09-949-016-130943/c
; Sequence 130943, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 130943
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-130943

Query Match 1.7%; Score 41.4; DB 3; Length 601;
Best Local Similarity 56.1%; Pred. No. 0.14; Mismatches 61; Indels 0; Gaps 0;
Matches 78; Conservative 0

QY 1044 TGTGTAAGTAGTCCCGAGACAGGTGAAAGCTGTAAGCTGTAACCTTTGATAAGGAAA 1103
Db 253 TTTTAAAGTAGTCCCAAAAGGAGGAGAACTTAAAGCAATTCCTTAGTAAGGATAA 194
QY 1104 CTGTTGAAGACCTTATGCTAGTGTAAAGAAATTTGCAAGAGAGCGGAAAGATCA 1163
Db 193 AGTTTGATATCTTCAGTACAAATTTGTAACAATTTTTCATAATAAGGAAAAGGCTTA 134
QY 1164 ATGATCAGGAGTATGTAAA 1182
Db 133 ATTTTCAGGTAGACTGAAA 115

RESULT 12
US-09-949-016-15432/c
; Sequence 15432, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15432
; LENGTH: 67899

TYPE: DNA
ORGANISM: Human
US-09-949-016-15432

Query Match 1.7%; Score 41.4; DB 3; Length 67899;
Best Local Similarity 56.1%; Pred. No. 2;
Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 1044 TGTGTAAGTAGTCCCGAGACAGGTGAAAGCTGTAAGCTGTAACCTTTGATAAGGAAA 1103
Db 26889 TTTTAAAGTAGTCCCAAAAGGAGGAGAACTTAAAGCAATTCCTTAGTAAGGATAA 26830
QY 1104 CTGTTGAAGACCTTATGCTAGTGTAAAGAAATTTGCAAGAGAGCGGAAAGATCA 1163
Db 26829 AGTTTGATATCTTCAGTACAAATTTGTAACAATTTTTCATAATAAGGAAAAGGCTTA 26770
QY 1164 ATGATCAGGAGTATGTAAA 1182
Db 26769 ATTTTCAGGTAGACTGAAA 26751

RESULT 13
US-09-949-016-11870/c
; Sequence 11870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11870
; LENGTH: 67902
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11870

Query Match 1.7%; Score 41.4; DB 3; Length 67902;
Best Local Similarity 56.1%; Pred. No. 2;
Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 1044 TGTGTAAGTAGTCCCGAGACAGGTGAAAGCTGTAAGCTGTAACCTTTGATAAGGAAA 1103
Db 26889 TTTTAAAGTAGTCCCAAAAGGAGGAGAACTTAAAGCAATTCCTTAGTAAGGATAA 26830
QY 1104 CTGTTGAAGACCTTATGCTAGTGTAAAGAAATTTGCAAGAGAGCGGAAAGATCA 1163
Db 26829 AGTTTGATATCTTCAGTACAAATTTGTAACAATTTTTCATAATAAGGAAAAGGCTTA 26770
QY 1164 ATGATCAGGAGTATGTAAA 1182
Db 26769 ATTTTCAGGTAGACTGAAA 26751

RESULT 14
US-09-621-976-2813/c
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
; US-09-621-976-2813

Query Match 1.6%; Score 40; DB 3; Length 832;
Best Local Similarity 8.6%; Pred. No. 0.44;
Matches 28; Conservative 159; Mismatches 139; Indels 0; Gaps 0;
QY 1904 ATCTTGTGTTGGTTCGGATATCTAACCGGAGTGGACATTCCTAGTATTACATTTTCATGCC 1963
DB 329 WKTACASRYRYKTYGWWYWKMMSTMSTWYCMKCKMTRGRRCATWTTWARGMWSY 270
QY 1964 ACATTTCTTATGTTGTTGTTGTTGTTATTCACAAAGGTATGGGAAAGCTGAGGAATTT 2023
DB 269 ANGKWSMESANSMTCTMYKKGSTYVMKCTCATWCYWKYKWMWSKTCWSGSGGY 210
QY 2024 CTGCTGAACGATTCAGACATAGATGCGGCAATCCCTAACGAAACAAACACTGATTTCAAG 2083
DB 209 MTSYTSRSYSYMASWMTYMCWGRMWSYTYWMAWGKWKRYATTWRRAMMMWWAAWT 150
QY 2084 TAACTCTAGTAGACATCTTTTACACAACTAGTGAATCAAGATTAGTGGTTTGAT 2143
DB 149 MWYMMWMAWMSRGAATYRRMMMGYRYWWRKSYRRRCAYAWATKRSYWCWRW 90
QY 2144 TAGGGAATTTAAAGATGATTTCTTTTTCACAGATTATCCCATTCAGTGGAGGCC 2203
DB 89 KWKRCMMWMAWYKTMWBRACWKTRYWRWAWAWWMMWMTMMWYMYWWRKRRM 30
QY 2204 TAGAAATCTGTAGCGATCAGTTTG 2229
DB 29 MRKRSWSWMMWMAWGWTRWAARMWR 4

RESULT 15
US-09-949-016-88266/c
; Sequence 88266, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88266
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-88266

Query Match 1.6%; Score 39.6; DB 3; Length 601;
Best Local Similarity 49.1%; Pred. No. 0.48;
Matches 105; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
QY 791 GAAGTTCTCTGACATGACATGATGTCATGGGTGTCCTTTTAACTACAATTCGA 850

Db 246 GAATTCCTGTTCCAGATAAACAATTTATTATTATTACATTTATTATTATAATAATAACAT 187
QY 851 TTCTTTTGACTACTGATAGTCCTGTCAATTGAAGCTGTTTACACTGCTCTTTAAAGAAAGCTGA 910
DB 186 TATTATTAAATGTTATTACATTTATTAGGTAATTACCTGACAGTGTCTTTAAGTGGCAGA 127
QY 911 GCTTCGTTTCTACTGATCTTCTGCCATATTTGGGAGCGCAAGTTTCTCTGTTGTTTCTGTCG 970
DB 126 TACTACTATGTCATGATTTACATGGGTCTACTGAAACAAGTTTTTTTTTTTTTTTTTTTGT 67
QY 971 TTTGTTGATTTGTGGAAACAATTTGATTTCTTCTT 1004
DB 66 TTTTGTGTTTTTTTTTTGGAGACGAGTCTCGCT 33

Search completed: January 13, 2006, 07:19:43
Job time : 438 secs

This Page Blank (uspto)

[illegible]

1381	ACTGGATCTCTCTCCTCCTTGCGACACTTTATCTCTCTAAGTAAGGTACCTTAAATGTATCTTCTC	1444
Qy	TACTTTTGCTATGCTAGAGAAATTTACTTTGGATGGGAGCTTCTCTGTGTTCTCAATTTACCTCTTT	1500
Db	TACTTTTGCTATGCTAGAGAAATTTACTTTGGATGGGAGCTTCTCTGTGTTCTCAATTTACCTCTTT	1500
Qy	CAAAATTCCTATGTTTCATAGAACTCATCTGCAATTTAGGAAAGCACAAAGAAAGTAAGTACACA	1560
Db	CAAAATTCCTATGTTTCATAGAACTCATCTGCAATTTAGGAAAGCACAAAGAAAGTAAGTACACA	1560
Qy	GAGTGTTAGAAGGAAGAAACCCGGCTTTTCAGGAGTATAAAGAGTGTGAAGTACATCACTC	1620
Db	GAGTGTTAGAAGGAAGAAACCCGGCTTTTCAGGAGTATAAAGAGTGTGAAGTACATCACTC	1620
Qy	GTTCGTATAAACGAGTCAATCGTCTCTATCCTCATCTCTCTGTAAGCAATCAAGCTCATC	1680
Db	GTTCGTATAAACGAGTCAATCGTCTCTATCCTCATCTCTCTGTAAGCAATCAAGCTCATC	1680
Qy	TCTCTAATTAATTCATGAACCTAAATTTTCTGATTTGATTTGTTTCTCTGTTAGGTTCTTGATAA	1740
Db	TCTCTAATTAATTCATGAACCTAAATTTTCTGATTTGATTTGTTTCTCTGTTAGGTTCTTGATAA	1740
Qy	GAAAGCTCAAGTTCCTGACATCTTCTCCGGAACCTATAAGGTCAATACCCGCAACAAGACA	1800
Db	GAAAGCTCAAGTTCCTGACATCTTCTCCGGAACCTATAAGGTCAATACCCGCAACAAGACA	1800
Qy	TTATGATTTTCAGTCTATAAACATCCTCATCGTCTCTTCCGAGGTACAGTTCCTTCTCTCTCTC	1860
Db	TTATGATTTTCAGTCTATAAACATCCTCATCGTCTCTTCCGAGGTACAGTTCCTTCTCTCTCTC	1860
Qy	GTCCATAGTATAACATAGGGGAGCTAATCGTTCCTCTTCAATGATCTTTCTGTGTTGTTTCGG	1920
Db	GTCCATAGTATAACATAGGGGAGCTAATCGTTCCTCTTCAATGATCTTTCTGTGTTGTTTCGG	1920
Qy	ATATCTAACCCGAGTGGACATTTCTCTAGTATTACATTCATGCCACATTTCTTATGTGTTT	1980
Db	ATATCTAACCCGAGTGGACATTTCTCTAGTATTACATTCATGCCACATTTCTTATGTGTTT	1980
Qy	GTTTGTTGTTTATTCCAAAGGTATCGGAAAAAGCTGTAGAAATTTCTGCTGAAACGATTCGA	2040
Db	GTTTGTTGTTTATTCCAAAGGTATCGGAAAAAGCTGTAGAAATTTCTGCTGAAACGATTCGA	2040
Qy	CATAGATGGCGCAATCCCTTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAAC	2100
Db	CATAGATGGCGCAATCCCTTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAAC	2100
Qy	ATCTTTTGCACAAACCTACTCGAATCAAGATTAGTGGTTTTCATTTAGGGAATTTAAAAAGAT	2160
Db	ATCTTTTGCACAAACCTACTCGAATCAAGATTAGTGGTTTTCATTTAGGGAATTTAAAAAGAT	2160
Qy	GATTTTCTTTTTTTCACAGATTCATCCCATTCAGTGGAGGCCCTAGAAAAATGTGTAGCGC	2220
Db	GATTTTCTTTTTTTCACAGATTCATCCCATTCAGTGGAGGCCCTAGAAAAATGTGTAGCGC	2220
Qy	ATCAGTTTGCAATTTGATAGGAGCAATTTGTGGCACTCTCGGGTGTGTTTCTCAGCGGTTAAACG	2280
Db	ATCAGTTTGCAATTTGATAGGAGCAATTTGTGGCACTCTCGGGTGTGTTTCTCAGCGGTTAAACG	2280
Qy	TTGAGCTGGTTCCTGATCAGACCACTTAGCATGACCAAGGAGCAACCATATACACACCACA	2340
Db	TTGAGCTGGTTCCTGATCAGACCACTTAGCATGACCAAGGAGCAACCATATACACACCACA	2340
Qy	ATGTATGCCAATGTTTCTCACACTCGAGAGATTAATATGAGAGTGTCTGTGTTTGTGTAGAAATG	2400
Db	ATGTATGCCAATGTTTCTCACACTCGAGAGATTAATATGAGAGTGTCTGTGTTTGTGTAGAAATG	2400
Qy	ATTCCAAATTTCTTAATCTGTATATTTTCAATTTTCAGGGAATGTATATGAAGGTGAGCCA	2460
Db	ATTCCAAATTTCTTAATCTGTATATTTTCAATTTTCAGGGAATGTATATGAAGGTGAGCCA	2460
Qy	AAAGTAA 2467	
Db	AAAGTAA 2467	

		Matches 2466; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Qy	1801	TTATGATTTAGTCTATTAACATCCATCTTCTCGAGGTACAGTTCCTTCTCTCTC	1860
Db	2810	TTATGATTTAGTCTATTAACATCCATCTTCTCGAGGTACAGTTCCTTCTCTCTC	2869
Qy	1861	GTCCATAGTATAACATAGGGAGCTATCTCTCTTCAATGATCTTTGTGTGTTGG	1920
Db	2870	GTCCATAGTATAACATAGGGAGCTATCTCTCTTCAATGATCTTTGTGTGTTGG	2929
Qy	1921	ATATCTAACCGGAGTGGACATTCCTAGTATTACATTTCAATGATCTTTGTGTGTT	1980
Db	2930	ATATCTAACCGGAGTGGACATTCCTAGTATTACATTTCAATGATCTTTGTGTGTT	2989
Qy	1981	GTGTGTTGTTATTCCTCAAGGTATGGGAAAGCTGAGGAAATTTCTGCTGAACGATTCGA	2040
Db	2990	GTGTGTTGTTATTCCTCAAGGTATGGGAAAGCTGAGGAAATTTCTGCTGAACGATTCGA	3049
Qy	2041	CATAGATGGCGAATCCCTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC	2100
Db	3050	CATAGATGGCGAATCCCTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC	3109
Qy	2101	ATCTTTTGACACAAACTACTGAACTCAAGATTAGTGGTTTGAATAGGGAATTTAAAGAT	2160
Db	3110	ATCTTTTGACACAAACTACTGAACTCAAGATTAGTGGTTTGAATAGGGAATTTAAAGAT	3169
Qy	2161	GATTTTCTTTTTCACAGATTCATCCATTCAGTGGAGGGCTAGAAAAATGTGTAGGG	2220
Db	3170	GATTTTCTTTTTCACAGATTCATCCATTCAGTGGAGGGCTAGAAAAATGTGTAGGG	3229
Qy	2221	ATCAGTTTGCATTTAGTGGGAAATTTGGCACTCGCGTGTCTTCTCAGCGGTTAAACG	2280
Db	3230	ATCAGTTTGCATTTAGTGGGAAATTTGGCACTCGCGTGTCTTCTCAGCGGTTAAACG	3289
Qy	2281	TTGAGCTGTTCTCGATCAGACCATTAGCATGACACACAGGAGCAACCATACACACCA	2340
Db	3290	TTGAGCTGTTCTCGATCAGACCATTAGCATGACACACAGGAGCAACCATACACACCA	3349
Qy	2341	ATGTATGCCAATGTTCTCACCTCGAGAGATTAATGAGAGTGTCTGTTTGTAGATG	2400
Db	3350	ATGTATGCCAATGTTCTCACCTCGAGAGATTAATGAGAGTGTCTGTTTGTAGATG	3409
Qy	2401	ATTCAAATTTCTTAATGCTGATATTTCAATTTCAAGGATTTGATGAGGCTGAGCCA	2460
Db	3410	ATTCAAATTTCTTAATGCTGATATTTCAATTTCAAGGATTTGATGAGGCTGAGCCA	3469
Qy	2461	AAAGTAA 2467	
Db	3470	AAAGTAA 3476	
RESULT 3			
US-10-751-235-7			
; Sequence 7, Application US/10751235			
; Publication No. US20050150002A1			
; GENERAL INFORMATION:			
; APPLICANT: DellaPenna, Dean			
; APPLICANT: Tian, Li			
; APPLICANT: Kim, Joonvul			
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid			
; FILE REFERENCE: MSU-08604			
; CURRENT APPLICATION NUMBER: US/10/751,235			
; CURRENT FILING DATE: 2004-01-02			
; NUMBER OF SEQ ID NOS: 74			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 7			
; LENGTH: 2467			
; TYPE: DNA			
; ORGANISM: Arabidopsis thaliana			
US-10-751-235-7			
Query Match 99.9%; Score 2465.4; DB 9; Length 2467;			
Best Local Similarity 100.0%; Pred. No. 0;			

1081 GCGTGTAACTTTGATTAAGGGAACCTGTTGAAGACCTTATGCTTAAGTGTAAAGAAATTGT 1140
1081 GCGTGTAACTTTGATTAAGGGAACCTGTTGAAGACCTTATGCTTAAGTGTAAAGAAATTGT 1140
1141 CGAAGAGAGAGCGGAAGAAATCAATGATGAGGAGTATGTAATGATGCTGACCCAGTAT 1200
1141 CGAAGAGAGAGCGGAAGAAATCAATGATGAGGAGTATGTAATGATGCTGACCCAGTAT 1200
1201 CCGGCTTTCTGCTTTCGCAAGCAGAGAGAGGTTTAAACTTTTCCCTTAACTTATTAAG 1260
1201 CCGGCTTTCTGCTTTCGCAAGCAGAGAGAGGTTTAAACTTTTCCCTTAACTTATTAAG 1260
1261 CAAATTTGGCCCTTCATTTACGATTAATCGAAGCTGATGTTGCAATTTGAGGCTTTTCCG 1320
1261 CAAATTTGGCCCTTCATTTACGATTAATCGAAGCTGATGTTGCAATTTGAGGCTTTTCCG 1320
1321 GTATCAAGTGTGAGTTACGGGATGATCTTCTCAATGCTGTAAGGCTCATGAAC 1380
1321 GTATCAAGTGTGAGTTACGGGATGATCTTCTCAATGCTGTAAGGCTCATGAAC 1380
1381 ACTGGATCTGCTCCTCACTTGGACACTTTATCTCTAAAGTAAGGTACCTTAAATGTATCTTC 1440
1381 ACTGGATCTGCTCCTCACTTGGACACTTTATCTCTAAAGTAAGGTACCTTAAATGTATCTTC 1440
1441 TACTTGTCTATGCTAGAGAAATTAATCTGGAGCTTCTCTGATGGAGCTTCTCTGTTCTCAATTA 1500
1441 TACTTGTCTATGCTAGAGAAATTAATCTGGAGCTTCTCTGATGGAGCTTCTCTGTTCTCAATTA 1500
1501 CAAATCTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAGCAACAGAGAGTAGACA 1560
1501 CAAATCTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAGCAACAGAGAGTAGACA 1560
1561 GAGTGTAGAGAGAGAAACCGGCTTTCGAGATATAAGGAGTTGAAGTACATCACTC 1620
1561 GAGTGTAGAGAGAGAAACCGGCTTTCGAGATATAAGGAGTTGAAGTACATCACTC 1620
1621 GTTGTATAAACGAGTCAATGCGTCTCTATCTCTCATCTCTCTGTAAGCAATCAAGCTCATC 1680
1621 GTTGTATAAACGAGTCAATGCGTCTCTATCTCTCATCTCTCTGTAAGCAATCAAGCTCATC 1680
1681 TCTCTAATTAATCATGAACTAAATTTCTGATTTGATTTGTTCTGTTAGGCTTTGATAA 1740
1681 TCTCTAATTAATCATGAACTAAATTTCTGATTTGATTTGTTCTGTTAGGCTTTGATAA 1740
1741 GAAGAGCTCAAGTCTCTGACATCTTCTCTGGGAACTATAAGGTCAATACCGGACAGACA 1800
1741 GAAGAGCTCAAGTCTCTGACATCTTCTCTGGGAACTATAAGGTCAATACCGGACAGACA 1800
1801 TTATGATTTCACTCTATAACATCCATCGTCTTCTCCGAGGTACAGTTCTCTCTCTCTC 1860
1801 TTATGATTTCACTCTATAACATCCATCGTCTTCTCCGAGGTACAGTTCTCTCTCTCTC 1860
1861 GTCCATAGTATAAACATAGGGAGCTTAATCTCTCTCTCAATGATCTTGTGTTGG 1920
1861 GTCCATAGTATAAACATAGGGAGCTTAATCTCTCTCTCAATGATCTTGTGTTGG 1920
1921 ATATCTAACCGAGTGGACATTTCTAGTATTAACATTCATGCCACATTTCTTATGTGTT 1980
1921 ATATCTAACCGAGTGGACATTTCTAGTATTAACATTCATGCCACATTTCTTATGTGTT 1980
1981 GTTGTGTTGTTATTCGAAAGGTATGGGAAAAAGCTGAGGAATTTCTGCTGAAACGATTCGA 2040
1981 GTTGTGTTGTTATTCGAAAGGTATGGGAAAAAGCTGAGGAATTTCTGCTGAAACGATTCGA 2040
2041 CATAGATGGCGAATCCCTTAACGAAACAAACATGATTTCAAGTAACTCAGTAGAACAC 2100
2041 CATAGATGGCGAATCCCTTAACGAAACAAACATGATTTCAAGTAACTCAGTAGAACAC 2100
2101 ATCTTTTGCACAACTCAAGTATGAGTGTGTTTGAATAGGGAAATTTAAAGAT 2160
2101 ATCTTTTGCACAACTCAAGTATGAGTGTGTTTGAATAGGGAAATTTAAAGAT 2160

2161 GATTTCTTTTTCACAGATTCATCCATTCAGTGGAGGCTAGAAAATGTGTAGCG 2220
2161 GATTTCTTTTTCACAGATTCATCCATTCAGTGGAGGCTAGAAAATGTGTAGCG 2220
2221 ATCAGTTTGCATTTGATGAGGCAATTTGGGCACTCGCGTGTCTTTCAGCGGTTAAACG 2280
2221 ATCAGTTTGCATTTGATGAGGCAATTTGGGCACTCGCGTGTCTTTCAGCGGTTAAACG 2280
2281 TTGAGCTGGTTCCTGATCAGACATTAGCATCAACAGGAGAACCATACACACCA 2340
2281 TTGAGCTGGTTCCTGATCAGACATTAGCATCAACAGGAGAACCATACACACCA 2340
2341 ATGATGCGAATGTTCTCAGCTCGAGAGTAAATGAGAGTCTGTTTGTGTTAGATG 2400
2341 ATGATGCGAATGTTCTCAGCTCGAGAGTAAATGAGAGTCTGTTTGTGTTAGATG 2400
2401 ATTCCAAATTTCTTAATGCTGATATTTCAATTTTCAGGAGTGTATATGAAGGTAGC 2460
2401 ATTCCAAATTTCTTAATGCTGATATTTCAATTTTCAGGAGTGTATATGAAGGTAGC 2460
2461 AAGGTAA 2467
2461 AAGGTAA 2467

RESULT 4

US-10-739-930-773
; Sequence 773, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 773
; LENGTH: 3377
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER3001_1
US-10-739-930-773

Query Match 32.5%; Score 803; DB 8; Length 3377;
Best Local Similarity 73.2%; Pred. No. 4.6e-208;
Matches 1347; Conservative 0; Mismatches 0; Indels 494; Gaps 5;

QY 1 ATGGAGTCTTCACTCTTTTCTCCATCTTCTCTTCTTACTCTTCTCTTCTCTTCTCTCTGCAAAA 60
DB 79 ATGGAGTCTTCACTCTTTTCTCCATCTTCTCTTCTTCTCTTCTCTTCTCTTCTCTTCTCTGCAAAA 138
QY 61 CTAGCGCTTTTATACACCAAAACCCAAATTCATCTCCATCAGATCTTCCATTTAG 120
DB 139 CTAGCGCTTTTATACCAAAACCCAAATTCATCTCCATCAGATCTTCCATTTAG 198
QY 121 AAACCCAAACCCAAACCTCGAGACCAATTCATCGAAATCCCAATCATGGTCAAGTCCCGAT 180
DB 199 AAACCCAAACCCAAACCTCGAGACCAATTCATCGAAATCCCAATCATGGTCAAGTCCCGAT 258
QY 181 TGGCTCACACACTCACTCGTACCCCTTTCTCAGGAAAAACGACGAGTCAGGTATACCA 240
DB 259 TGGCTCACACACTCACTCGTACCCCTTTCTCAGGAAAAACGACGAGTCAGGTATACCA 318
QY 241 ATCGGAGCGGAGCTCGAGATCTCGTGTCTCTCGGAGTGTCTCTCTTCTTACCT 300
DB 319 ATCGGAGCGGAGCTCGAGATCTCGTGTCTCTCGGAGTGTCTCTCTTCTTACCT 378
QY 301 CTCTACAAATGATGAATGAGTACGAGACCCATTTACCGTCTCGTGTGCTCTCTGTAAT 360
DB 379 CTCTACAAATGATGAATGAGTACGAGACCCATTTACCGTCTCTCTGCTGCTCTCTGTAAT 438

QY 887 TTACTGCTCTTAAGAGCTGAGCTTGGTTCTTACTGATCTTCTGCCATATTGGAAGC 946
| | | | |
Db 634 TTATATGCACTCAAGAGCAGAGCTTGGTTCTACAGATCTTTTGGCCATCTGGAAGT 693
| | | | |
QY 947 AAGTTTCTGCTGTTTTTCTGTTGTTTGTGATTGTTGGAACAATTGGATTCTTTGTTAA 1006
| | | | |
Db 694 TGGTTTC----- 700
| | | | |
QY 1007 TTGAGAGGTTGGTTGTTTTTTTTCAGATCGATGCAATTTGTGTAAGATAGTCCCGAGACAG 1066
| | | | |
Db 701 -----TTGTGCAAGATAATCCCAAGACAG 724
| | | | |
QY 1067 GTGAAGCTGAAAGCGCTGAATTTGATAAGGGAACCTGTTGAAGACCTTATTGCTAAG 1126
| | | | |
Db 725 ATAAAGCAGAGAATGCGGTTACGATTATAAGGAACACTGTTGAAGAGCTGATTATGAAG 784
| | | | |
QY 1127 TGTAAAGAAATTTGTCCAAAGAGAGCGGCAAGAAATCAATGATGAGGAGTATGTAATGAT 1186
| | | | |
Db 785 TGTAAAGAAATGTTGAAGCTGAAATGAACAGATTGAGGGTGAGGAATATGTAACGNA 844
| | | | |
QY 1187 GCTGACCAAGTATCTCGGTTTCTTGTGTCGAAGCAGAGAGAGGT 1233
| | | | |
Db 845 GGGATCCTAGCATTTCTACGCTTCTCTACTTGTAGCCGAGATGAGGT 891
| | | | |

RESULT 6

US-10-425-115-148190
; Sequence 148190, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 148190
; LENGTH: 2537
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_66678C.1
US-10-425-115-148190

Query Match 7.9%; Score 195.8; DB 8; Length 2537;
Best Local Similarity 63.2%; Pred. No. 5.3e-42;
Matches 371; Conservative 0; Mismatches 127; Indels 89; Gaps 1;
QY 647 GCGAGGCGTAGAGCGGTTGCCATCGCTTCAAGAGGATTTGTCTGTGATTGTGGA 706
| | | | |
Db 647 GGTGAGCGCAGAGCTGTTGGCTTCTCCCTGCAAAAGATTCTGTCCATTTATGCTGGA 706
| | | | |
QY 707 GAGAGTATTTCGCAATGTCAGAGAGGCTTGTGGAAGTTGCGAGCTTATGCAAGA 766
| | | | |
Db 707 CAAGTATTTGTAAATGTCTGAGAGACTAATAGACAACTTGAGCCATATGCTTTGAG 766
| | | | |
QY 767 CGAAGTCTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGTATGTCATTGGGTT 826
| | | | |
Db 767 TGGGAACTCTGCAATATGGAAGCGAGGTTTCTCAGTTGCAATTTGATGTTGTT 826
| | | | |
QY 827 GTCTCTTTTAACTACAAATTCGATTCTTTGACTACTGATGATGCTGTCATTGAGCTGT 886
| | | | |
Db 827 ATCATTTGTCACTACAAATTTGATTTCCCTCACAACAGATAGTCTGTCAATTGATGCTGT 886
| | | | |
QY 887 TTACTGCTCTTAAAGAGCTGAGCTTCTGTTCTTACTGATCTTCTGCCATATTGGAAGGC 946
| | | | |
Db 887 TTACTGCACTCAAGAGAGCAGAGCTTCTGTTCTACAGATCTTTTGGCCATCTACTGGAAGGT 946
| | | | |
QY 947 AAGTTTCTGCTGTTTTTTTCTGTTGTTGTTGATTGTTGGAACAATTTGGATTCTTTGTTAA 1006
| | | | |

Db 947 TGGTTTC----- 953
| | | | |
QY 1007 TTGAGAGGTTGGTTGTTTTTTTTCAGATCGATGCAATTTGTGTAAGATAGTCCCGAGACAG 1066
| | | | |
Db 954 -----TTGTGCAAGATAATCCCAAGACAG 977
| | | | |
QY 1067 GTGAAGCTGAAAGCGCTGAATTTGATAAGGGAACCTGTTGAAGACCTTATTGCTAAG 1126
| | | | |
Db 978 ATAAAGCAGAGAATGCGGTTACGATTATAAGGAACACTGTTGAAGAGCTGATTATGAAG 1037
| | | | |
QY 1127 TGTAAAGAAATTTGTCCAAAGAGAGCGGCAAGAAATCAATGATGAGGAGTATGTAATGAT 1186
| | | | |
Db 1038 TGTAAAGAAATGATGGAAGCTGAAATGAACAGATTGAGGGTGAGGAATATGTAACGAA 1097
| | | | |
QY 1187 GCTGACCAAGTATCTCGGTTTCTTGTGTCGAAGCAGAGAGAGGT 1233
| | | | |
Db 1098 GGGATCCTAGCATTTCTACGCTTCTCTACTTGTAGCCGAGATGAGGT 1144
| | | | |

RESULT 7

US-10-751-235-24
; Sequence 24, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-751-235-24

Query Match 7.6%; Score 187.2; DB 9; Length 1086;
Best Local Similarity 61.9%; Pred. No. 7.2e-40;
Matches 369; Conservative 0; Mismatches 138; Indels 89; Gaps 1;
QY 647 GCGAGGCGTAGAGCGGTTGCCATCGCTTCAAGAGGATTTGTCTGTGATTGTGGA 706
| | | | |
Db 489 GGTGAGCGTAGAGCGGTTGTACCATCTCTACACAAAGATTCTCTCAGTAATGCTCGA 548
| | | | |
QY 707 GAGAGTATTTCGCAATGTCAGAGAGGCTTGTGGAAGTTGCGAGCTTATGCAAGA 766
| | | | |
Db 549 TAAAGTGTCTGTAATGTGCTGAGAGATTGTTGGAAGGCTCGAGACTTATGCTTTGAG 608
| | | | |
QY 767 CGAAGTCTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGTATGATTTGGGTT 826
| | | | |
Db 609 TGGTGAACCTGTTAATATGGAAGCGAGGTTTCTCAAATGACATTAGATGATTTGGTTT 668
| | | | |
QY 827 GTCTCTTTTAACTACAAATTTGATTTCTTGCATCTGATAGTCTGTCAATTGAAGCTGT 886
| | | | |
Db 669 ATCTTGTTCACCTACAACTTTGATTCCCTCACATCAGATAGTCTGTTATTGATGCTGT 728
| | | | |
QY 887 TTACTGCTCTTAAAGAGCTGAGCTTCTGTTCTACTGATCTTCTGCCATATTGGAAGGC 946
| | | | |
Db 729 TTACTGCACTCAAGAGAGCTGAGGCTGTTCTACAGATCTTTTACCATACTCG- 783
| | | | |
QY 947 AAGTTTCTGCTGTTTTTTTCTGTTGTTTGTGTTGGAACAATTTGGATTCTTTGTTAA 1006
| | | | |
Db 784 ----- 783
| | | | |
QY 1007 TTGAGAGGTTGGTTGTTTTTTTTCAGATCGATGCAATTTGTGTAAGATAGTCCCGAGACAG 1066
| | | | |
Db 784 -----CAGATCGATTGCTGTGCAAGATTGTTCTTAGACAG 819
| | | | |

Qy 1067 GTGAAGCTGAAAGGCTGTAACTTTGATAGGGAAAGCTTTGAAGACCTTTATTGCTAAG 1126
Db |||||
820 ATAAAGCGGAAAAGCAGTTAAACAAATAAGGAATACCGTTGAAGAGCTAATTACAAAA 879
Qy 1127 TGTAAAGAAATGTCGAAAGAGAGGCGGAAAGAAATCAATGATGAGGAGTATGTAATGAT 1186
Db |||||
880 TGCAGGCATCTGTAGATGCTGAAATGAACAGATTGAGGGTGAAGAATAATGTAATGAG 939
Qy 1187 GCTGACCAAGTATCTCGGTTCTTGTGCAAGCAGAGAGAGGTTTAAACTTT 1242
Db |||||
940 GCAGATCCTAGCATCTCGGTTTACTTGTAGCCGTGAAGAGGTGAGCAGTTT 995

RESULT 8
US-10-751-235-23
; Sequence 23, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Hordeum vulgare
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1587)..(1587)
; OTHER INFORMATION: n is a, c, g, or t

US-10-751-235-23

Query Match 7.6%; Score 187.2; DB 9; Length 1638;
Best Local Similarity 61.9%; Pred. No. 9.2e-40;
Matches 369; Conservative 0; Mismatches 138; Indels 89; Gaps 1;

Qy 647 GCGAGGCGTAGAGCGGTGGTCCATCGCTTACAGAGGATTTGTCTGATGATGTGA 706
Db |||||
474 GGTGAGCGTAGAGCAGTTGTACCATCTCTACAAAAGATTCTCTCAGTAATGTTGA 533
Qy 707 GAGAGTATCTCAAAATGTCAGAGAGGCTTGTGAGAAGTTGCAGCCTTATGCGAAGA 766
Db |||||
534 TAAAGTGTGTAATGCTGTGAGAGATTGTGGAAGCTCGAGACATATGCTTTGAG 593
Qy 767 CGGAAGTGTGTAATGCGAAGCGAAGTTCTCTCAGATGACACTTGTGATGTCATTTGGGTT 826
Db |||||
594 CGGTGAACCTGTTAATATGGAAGCGAGATTTCTCAAAATGACACTAGATGATGTTGGTTT 653
Qy 827 GTCTCTTTTAACTACAATTCGATTTCTTTGACTACTGATGATGCTCTGCTCATTTGAAGCTGT 886
Db |||||
654 GTCTTTGTTCAACTACAACCTTTGATTCCTCAATCAGATGCTCTGTTATTGATGCTGT 713
Qy 887 TTACACTGCTCTTAAAGAGCTGAGCTGCTTCTACTGATCTTCTGCCATATTTGAAGGC 946
Db |||||
714 TTACACCGCACTGAAAGAGCAGAGGCTGTTCTACAGATCTTTTACCATCTGG----- 768
Qy 947 AAGTTTCTGTGTTTTTTTCTGTGGTTTGTGATGTTGTGGAACAATTTGGATTCTTTTAA 1006
Db |||||
769 ----- 768

Qy 1007 TTGAGAGGGTTGGTTGTTTTTTTTCAGATCGATGCAATTTGTGAAGATAGTCCCGACAG 1066
Db |||||
769 -----CAGATTGATTTGCTGTGCAAGATTGTTCTCAGACAG 804

Qy 1067 GTGAAGCTGAAAGGCTGTAACTTTGATAGGGAAAGCTTTGAAGACCTTTATTGCTAAG 1126
Db |||||
805 ATCAAAGCAGAAAGGAGTAAACAAATAAGGAATACTGTTGAGAGCTTAATTATAAAA 864

Qy 1127 TGTAAAGAAATGTCGAAAGAGAGGCGGAAAGAAATCAATGATGAGGAGTATGTAATGAT 1186
Db |||||
865 TGCAGGCATCTGTAGATGCTGAAATGAACAGATTGAGGGTGAAGAATAATGTAATGAG 924
Qy 1187 GCTGACCAAGTATCTCGGTTCTTGTGCAAGCAGAGAGAGGTTTAAACTTT 1242
Db |||||
925 GCAGATCCTAGCATCTCGGTTTACTTGTAGCCGTGAAGAGGTGAGCAGTTT 980

RESULT 9
US-10-424-599-122968
; Sequence 122968, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 122968
; LENGTH: 667
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_82046C.1

US-10-424-599-122968

Query Match 7.5%; Score 185; DB 7; Length 667;
Best Local Similarity 69.9%; Pred. No. 2.2e-39;
Matches 283; Conservative 0; Mismatches 110; Indels 12; Gaps 2;

Qy 98 TCTCCATCAGATCTCTCCATTCGAGAAACCCAAACCCAAACTCGAGACCAATTCATCGAAAT 157
Db |||||
150 TCTCGTCAATCTCTCCATCAACAAACCC-----CCTGAGCGCGGCACCAACAGC 200
Qy 158 CCCAATCAATGGGTGAGTCCCGATTGGCTCAACAACAATCTACTGCTGATCTCC 217
Db |||||
201 CCAGCTCTGGGTGAGTCCCGGCTGAGCCCGGCTGCTCCCTCTCTCCGCTCCCTCACCGCGGCA 260
Qy 218 AAAACGACGAGTCAAGTATACCAATCGGAGCGGAGCTCGACGATGCTGCTGATCTCC 277
Db |||||
261 ---ACGACGCTCTCCGGCATCCCGCTCGCCCAAGCTCGACGAGTCTCGATCTCC 317
Qy 278 TCGGAGGTGCTCTCTTTTACCTCTTACAAATGGATGAATGAGTACGAGCCCATTTACC 337
Db |||||
318 TCGGCGGCGCTCTCTCTCTCCGCTCTTCAAGTGGATGAGGACTAGCGGCCCATCTACC 377
Qy 338 GTCTCGCTGCTGCTCTCTGTAATTTGTAATTTGTGAGCGACCCAGGATAGCTAAACATG 397
Db |||||
378 GCCTGCGCGCGCGCCCGGAAACTTCGTTGTGTTAGCGACCTGCGCATTTGCCAAGCAG 437
Qy 398 TTTTGAAGAAATTCACAAAGTACGCTTAAGGCTTAGTCGCTGAAGTCTCTGAATTTCTAT 457
Db |||||
438 TGTCTAGGAACACTACGCTAAATATGCTAAAGGCTCGTGTCTGAGGCTCTCGAGTCTCTGT 497
Qy 458 TTGCTTTCGGTTTTCGCTATCGCTGAAGGACCTCTTTTGGACAGTAA 502
Db |||||
498 TCGGCTCGGTTTTCGCCATCGCGGAAGGTCCTCTCTCGAGCGCAA 542

RESULT 10
US-10-751-235-22
; Sequence 22, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li

APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15200
; LENGTH: 1118
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21066C.1
US-10-437-963-15200

Query Match 7.4%; Score 183; DB 9; Length 1686;
Best Local Similarity 61.8%; Pred. No. 1.3e-38;
Matches 363; Conservative 0; Mismatches 135; Indels 89; Gaps 1;

QY 647 GCGAGGCTAGAGCGGTGCTCCATCGCTTACAGGAGGTATTTGCTGTGATGTGCA 706
DB 522 GGTGAGACGTGATCAGTTGTACCATCTCTACCAACAGATTCTCTCGGTGATGTGCA 581
QY 707 GAGAGTATTCTGCAATGTCGAGAGAGGCTTGTGAGAAAGTTGCGAGCTTATGCGAGAGCA 766
DB 582 CAGAGTTTTTGTAAATGTCTGAGAGATTAGTGGAGAGCTTGAGACATCTGCTTTAAG 641
QY 767 CGAAGTGTCTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGTATGTGCTGGTT 826
DB 642 TGGCAAACTGTAAATATGGAAGCAAGTTCTCTCAATGACTTTAGATGTGATGGTTT 701
QY 827 GTCCTTTTAACTCAATTCGATTTCTTGTACTACTGATGCTGCTGCTTCAATGAGCTGT 886
DB 702 GTCTTGTTCATTAACAATTTTGAATTCCTCAGATGAGCTGTTATGATGCTGT 761
QY 887 TTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGGC 946
DB 762 TTACACTGCACTCAAGAGAGCAAGTTCTGTTCTACAGATCTTTTACCATCTGCA---- 817
QY 947 AAGTTTCTGTGTTTTCTGTGTTTTGTGATGTGTGGAACAATTTGGATTTCTTTTAA 1006
DB 818 ----- 817
QY 1007 TTGAGAGGTTTGGTTGTTTTTTTTCAGATCGATGATTTGTGTAGATAGTCCGAGACG 1066
DB 818 -----AGATTGATTTGCTGTGTCGAAGATTGTTCTCTAGACAA 852
QY 1067 GTGAAGCTGAAAGCGCTTAATTTGATAAGGGAACCTGTTGAAGACCTTATTTGTAAG 1126
DB 853 ATAAAGCGAAAGGCGAGTTAAATCATCAGAAATACGTTGAGGACCTAATTTACCAA 912
QY 1127 TGTAAAGAAATTTGTGAAAGAGAGCGGAAAGAAATCAATGATGAGGAGTATGTAATGAT 1186
DB 913 TCGAAGAAGATTGTAGTCTGAGAATGAACAAATTTGAGGTTGAGGAATATGTAATGAG 972
QY 1187 GTTGACCCAGTATCTGCTGCTTCTGCTTGAAGCAGAGAGAGT 1233
DB 973 GCAGACCTTAGCATCTCGCATTCCTACTTGTGTCGCTGGAAGAGT 1019

RESULT 11
US-10-437-963-15200
; Sequence 15200, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15200
; LENGTH: 1118
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21066C.1
US-10-437-963-15200

Query Match 7.3%; Score 181; DB 7; Length 1118;
Best Local Similarity 73.2%; Pred. No. 3.7e-38;
Matches 232; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 642 ACATAGGCGAGGCTAGAGCGGTGCTCCATCGCTTACAGGAGGTATTTGCTGTGATTT 701
DB 61 ACGAAGGTGAGACGTGATCAGTTGTACCATCTCTACACAAACGATTTCTCTCGGTGATG 120
QY 702 GTGAGAGAGATTTCTGCAAAATGTGAGAGAGGCTTGTGAGAAAGTTGCGAGCTTATGCA 761
DB 121 GTTGACAGAGTTTTTTGTAATGTCTGAGAGATTAGTGGAGAGCTTGAGACATCTGCT 180
QY 762 GAAGACGGAAGTGTCTGTAATATGGAAGCGAAGTTCTCTCAGATGACACTTGAATGCTATT 821
DB 181 TTAAGTGGCAACCTGTAAATATGGAAGCAAGTTCTCTCAATGACTTTTAGATGTGATTT 240
QY 822 GGGTTGTCTCTTTTAACTTACAAATTTGATTTCTTGTACTGATGATGCTGTCTGCTTAA 881
DB 241 GGTGTTCTCTGTTCAATTTGATTTCCCTCAGATGAGAGCTTGTGATGATTTATGAT 300
QY 882 GCTGTTTACACTGCTCTTAAAGAGCTGAGCTTGGTTCTTACTGATCTTCTGCCATATTTGG 941
DB 301 GCTGTTTACACTGCACTCAAGGAGCAGAACTTGGTTCTTACAGATCTTTTACCATCTGS 360
QY 942 AAGGCAAGTTTCTCTGTG 958
DB 361 AAGATTGATTTGCTGTG 377

RESULT 12
US-10-751-235-27
; Sequence 27, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Helianthus annuus
US-10-751-235-27

Query Match 6.9%; Score 169.6; DB 9; Length 624;
Best Local Similarity 69.6%; Pred. No. 3.4e-35;
Matches 245; Conservative 0; Mismatches 104; Indels 3; Gaps 1;
QY 608 TTCTCTATTTGGGCTTAGCTGAAGTCTGATTTTTCATAGGCGAGCGGTAGAGCGGTGCT 667
DB 117 TTCTGGTTTGGCAATTGCTGAAGGCTCTCTTTGGAC---TGCAAGCGCGAGGCTGTAGT 173
QY 668 TCCATCGTTTACAGAGGATTTTCTGTGTTTGTGAGAGATTTCTTGCANATGTGC 727

174 TCATCAGTTCACAAGAAGTACTTATCAGTAATATAGTTGATCGTGTTATTTTGCAATGCTC 233

728 AGAGAGGCTTGTGTGAGAAAGTTGCAGCCCTTATGCAAGAAGACGGAAGTCTGTGTAATATGGA 787

234 CGAGAGGCTTGTGTGAGAAAGTAAAGTATCATACGACGCAAGTGACACGCTCTGTTAAACATGGA 293

788 AGCGAAGTTCCTCTCAGATGACACATGTGATGTCATTTGGGTGTGTCTCTTTTAACTTACAATTT 847

294 GCAACAGATTTTCGCAAGTTAAACCTTGTATGTATTGCTCTAGCCGATATTTTAACTTACAATTT 353

848 CGATCTTTTGACTACTGATAGTCTCTGTCATTTGGAAGCTGTTTTACACTGCTCTTTAAAGAAGC 907

354 TGACTCACTTACGGCCGATAGTCTCTGTAATTTGAATCTGTTTATACCGCACTTAAAGAAGC 413

908 TGAGCTTCCTTCTACTCTGATCTCTTGCCATATTGGAAGCGCAAGTTTTCCTGTGT 959

414 TGAAGCCCGTTCAACTGATCTTTTGCCATATTGGAAGATTAAGTGGCGTTAATGT 465

RESULT 13

US-10-751-235-25

; Sequence 25, Application US/10751235

; Publication No. US20050150002A1

; GENERAL INFORMATION:

; APPLICANT: DellaPenna, Dean

; APPLICANT: Tian, Li

; APPLICANT: Kim, Joonyul

; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Ca

; TITLE OF INVENTION: Metabolism in Plants

; FILE REFERENCES: MSU-08604

; CURRENT APPLICATION NUMBER: US/10/751,235

; CURRENT FILING DATE: 2004-01-02

; NUMBER OF SEQ ID NOS: 74

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 25

; LENGTH: 839

; TYPE: DNA

; ORGANISM: Lycopersicon esculentum

US-10-751-235-25

```

332 TTGTAGAAACTGAGGCTGAGGAGTTAATGAAGTAGACTGCTGAATGATAGAGATCCAA 391
1197 GTATCTCTGCGTTTCTTGCTTCAAGCAGAGAGAGGTTTAAACTTTTTCCTT 1249
392 GCATCTCTCGATTTTGTCTTGCTAGCCGTGAGGAGGTTTCAAGTTTACAACTT 444

RESULT 14
US-10-751-235-26
; Sequence 26, Application US/10751235
; Publication No. US20050150002A1
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Tian, Li
; APPLICANT: Kim, Joonyul
; TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid
; TITLE OF INVENTION: Metabolism in Plants
; FILE REFERENCE: MSU-08604
; CURRENT APPLICATION NUMBER: US/10/751,235
; CURRENT FILING DATE: 2004-01-02
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Zea mays
US-10-751-235-26

```

RESULT 15
US-10-437-963-15203
; Sequence 15203, Application US/10437963

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 19:33:24 ; Search time 1445 Seconds
(without alignments)
11378.406 Million cell updates/sec

Title: US-10-751-235-5
Perfect score: 2467
Sequence: 1 atggagttctactcttttc.....tgaagtgagccaaaggttaa 2467

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_21.*
1: Geneseqm1980s.*
2: Geneseqm1990s.*
3: Geneseqm2000s.*
4: Geneseqm2001as.*
5: Geneseqm2001bs.*
6: Geneseqm2002as.*
7: Geneseqm2002bs.*
8: Geneseqm2003as.*
9: Geneseqm2003bs.*
10: Geneseqm2003cs.*
11: Geneseqm2003ds.*
12: Geneseqm2004as.*
13: Geneseqm2004bs.*
14: Geneseqm2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2467	100.0	2467	14 AEB16909	Aeb16909 Thale cre
2	2467	100.0	4170	14 AEB16910	Aeb16910 Thale cre
3	2465.4	99.9	2467	14 AEB16911	Aeb16911 Thale cre
4	803	32.5	3377	13 ADT15447	Adt15447 Plant cDN
5	195.8	7.9	1804	13 ADX61765	Adx61765 Plant ful
6	187.2	7.6	1086	14 AEB16926	Aeb16926 Wheat cyt
7	187.2	7.6	1638	14 AEB16925	Aeb16925 Barley cy
8	183	7.4	1686	14 AEB16924	Aeb16924 Rice cyto
9	169.6	6.9	624	14 AEB16929	Aeb16929 Sunflower
10	161	6.5	839	14 AEB16927	Aeb16927 Tomato cy
11	144.6	5.9	531	14 AEB16928	Aeb16928 Maize cyt
12	91	3.7	1788	14 AEB16943	Aeb16943 Thale cre
13	91	3.7	2057	14 AEB16942	Aeb16942 Thale cre
14	90.8	3.7	1278	14 AEB16947	Aeb16947 Soybean c
15	87.4	3.5	1078	13 ADO82389	Ado82389 Plant ful
16	84.6	3.4	1737	14 AEB16949	Aeb16949 Tomato cy
17	83.4	3.4	1899	14 AEB16945	Aeb16945 Rice cyt
18	76	3.1	2239	13 ADO81358	Ado81358 Plant ful
19	73.2	3.0	1731	14 AEB16957	Aeb16957 Soybean c

20	73.2	3.0	1831	2 AAX60780	Aax60780 Soybean c
21	72.8	3.0	1031	14 AEB16948	Aeb16948 Wheat cyt
22	72.8	3.0	1527	14 AEB16946	Aeb16946 Barley cy
23	72.8	3.0	1926	14 AEB16956	Aeb16956 Pea CYP97
24	72	2.9	1734	14 AEB16955	Aeb16955 Thale cre
25	68.8	2.8	1101	14 AEB16950	Aeb16950 Green alg
26	65.2	2.6	5071	14 AEB16944	Aeb16944 Rice cyto
27	59.4	2.4	579	10 ABX56831	Abx56831 Arabidops
28	57.8	2.3	404	3 AAA68063	Aaa68063 Pinus rad
29	57.8	2.3	404	10 ADD41813	Add41813 Cinnammat
30	56.6	2.3	1980	14 AEB16959	Aeb16959 Diatom cy
31	51.8	2.1	1299	13 ADS57092	Ads57092 Bacterial
32	51.2	2.1	1321	13 ADT44250	Adt44250 Bacterial
33	51	2.1	2000	8 ADA71938	Ada71938 Rice gene
34	50	2.0	297	12 ADN37071	Adn37071 Tobacco c
35	50	2.0	297	12 ADO85319	Ado85319 Nicotiana
36	50	2.0	313	12 ADN37107	Adn37107 Tobacco c
37	50	2.0	313	12 ADO85355	Ado85355 Nicotiana
38	50	2.0	1595	12 ADN10644	Adn10644 Nicotiana
39	50	2.0	1595	14 ADZ65332	Adz65332 Tobacco f
40	50	2.0	1595	14 ADZ65879	Adz65879 Tobacco c
41	49.2	2.0	2000	11 ACL35887	Acl35887 Rice stre
42	48.6	2.0	2000	11 ACL37108	Acl37108 Rice stre
43	47.8	1.9	259	12 ADN37087	Adn37087 Tobacco c
44	47.8	1.9	259	12 ADO85335	Ado85335 Nicotiana
45	47.6	1.9	2000	8 ADA71938	Ada71938 Rice gene

ALIGNMENTS

RESULT 1

AEB16909
ID AEB16909 standard; cDNA; 2467 BP.

AC AEB16909;

DT 08-SEP-2005 (first entry)

XX Thale cress LUT1 (cytochrome P450 97C1) cDNA, SEQ ID NO: 5.

XX Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;
XX cytochrome P450 97C1; gene; ss; chromosome 3.

OS Arabidopsis thaliana.

XX Key Location/Qualifiers
XX CDS 1..2467

FT /product= "Thale cress LUT1 protein"
FT /transl_except= (pos:496..650, aa:Thr-Ala)
FT /transl_except= (pos:942..1037, aa:Lys-Ile)
FT /transl_except= (pos:1242..1326, aa:Val-Ser)
FT /transl_except= (pos:1420..1523, aa:Lys-Asn)
FT /transl_except= (pos:1662..1841, aa:Val-Leu)
FT /transl_except= (pos:1839..2005, aa:Val-Trp)
FT /transl_except= (pos:2081..2183, aa:Lys-Phe)
FT /transl_except= (pos:2340..2440, aa:Asn-Gly)

US2005150002-A1.

07-JUL-2005.

02-JAN-2004; 2004US-00751235.

02-JAN-2004; 2004US-00751235.

(DELL/) DELLAPENNA D.

(TIAN/) TIAN L.

(KIMJ/) KIM J.

Dellapenna D, Tian L, Kim J;

XX

= instant application

Db 1610 TCCAACTCTTCTATTGGGCTTAGCTGAAGTCTGATTTTTCATAGCGGAGCGTAGAG 1669
Qy 661 CGGTGGTTCATCGCTTCAAGGAGGTATTTGTCTGTGATTTGTGGAGAGATATTCGCA 720
Db 1670 CGGTGGTTCATCGCTTCAAGGAGGTATTTGTCTGTGATTTGTGGAGAGATATTCGCA 1729
Qy 721 AATGTGAGAGAGGCTTGTGAGAGTTGCGAGCCTTATGCGAGACCGGAAGTCTGTGA 780
Db 1730 AATGTGAGAGAGGCTTGTGAGAGTTGCGAGCCTTATGCGAGACCGGAAGTCTGTGA 1789
Qy 781 ATATGGAAGCGAAGTCTCTCAGATGACACTTGATGTCATTTGGGTGTCTCTTTTAACT 840
Db 1790 ATATGGAAGCGAAGTCTCTCAGATGACACTTGATGTCATTTGGGTGTCTCTTTAACT 1849
Qy 841 ACAATTTCCGATTTCTTGACTACTGATGTCCTGTCATTTGAAGCTGTTTACACTGCTCTTA 900
Db 1850 ACAATTTCCGATTTCTTGACTACTGATGTCCTGTCATTTGAAGCTGTTTACACTGCTCTTA 1909
Qy 901 AAGAAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGCAAGTTTCCCTGTGT 960
Db 1910 AAGAAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGCAAGTTTCCCTGTGT 1969
Qy 961 TTTTCTGTGTTTGTGATTTGTGGAACAATTTGGAATCTTCTTTAACTGAGAGGTTTGG 1020
Db 1970 TTTTCTGTGTTTGTGATTTGTGGAACAATTTGGAATCTTCTTTAACTGAGAGGTTTGG 2029
Qy 1021 TTGTTTTTTTTCAGATCGATGTCATTTGTGAAGATGTCGCGAGACAGGTGAAGCTGAAAA 1080
Db 2030 TTGTTTTTTTTCAGATCGATGTCATTTGTGAAGATGTCGCGAGACAGGTGAAGCTGAAAA 2089
Qy 1081 GCGTGTAACTTTGATTAAGGAACTGTTGGAAGACCTTATTTGTAAGTGTGAAGAAATGT 1140
Db 2090 GCGTGTAACTTTGATTAAGGAACTGTTGGAAGACCTTATTTGTAAGTGTGAAGAAATGT 2149
Qy 1141 CGAAGAGAGCGGAAGAACTCAATGATGAGAGATGTAATGATGCTGACCCAAGTAT 1200
Db 2150 CGAAGAGAGCGGAAGAACTCAATGATGAGAGATGTAATGATGCTGACCCAAGTAT 2209
Qy 1201 CCGTGGTTCCTTGTGCAAGCAGAGAGAGGTTTAAACTTTTTCCTTTAAGTTTATAAG 1260
Db 2210 CCGTGGTTCCTTGTGCAAGCAGAGAGAGGTTTAAACTTTTTCCTTTAAGTTTATAAG 2269
Qy 1261 CAAATTTGGCCTTTCATTTATCGCATATCGAAGCTGATGTTGCATTTGTGAGGTTTTCAG 1320
Db 2270 CAAATTTGGCCTTTCATTTATCGCATATCGAAGCTGATGTTGCATTTGTGAGGTTTTCAG 2329
Qy 1321 GTATCAAGTGTGCGAGTTACGGGATGATCTTCTCAATGCTCGTAGCGGTCATGAAACC 1380
Db 2330 GTATCAAGTGTGCGAGTTACGGGATGATCTTCTCAATGCTCGTAGCGGTCATGAAACC 2389
Qy 1381 ACTGGATCTGCTCCTCACTTTGGAACCTTATCTCTTAAGTAAAGTACCTTAAATGATCTTC 1440
Db 2390 ACTGGATCTGCTCCTCACTTTGGAACCTTATCTCTTAAGTAAAGTACCTTAAATGATCTTC 2449
Qy 1441 TACTTTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTCTGTTCTCATTTTACCTCTT 1500
Db 2450 TACTTTGCTATGCTAGAGAAATTTACTTGGATGGAGCTTCTCTGTTCTCATTTTACCTCTT 2509
Qy 1501 CAAATTTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAAGCAACAAGAAAGTAGACA 1560
Db 2510 CAAATTTCTATGTTTCATAGAACTCATCTGCAATTAAGGAAAGCAACAAGAAAGTAGACA 2569
Qy 1561 GAGTGTTAGAAGGAAGAAACCCGGCTTTCGAGGATATAAGGATTAAGAGTTGAAGTACATCCTC 1620
Db 2570 GAGTGTTAGAAGGAAGAAACCCGGCTTTCGAGGATATAAGGATTAAGAGTTGAAGTACATCCTC 2629
Qy 1621 GTTGTATAAACGAGTCAATGCGTCTCTATCCTCATCTCTGTAAGCAATCAAGCTCATC 1680
Db 2630 GTTGTATAAACGAGTCAATGCGTCTCTATCCTCATCTCTGTAAGCAATCAAGCTCATC 2689
Qy 1681 TCTCTAATTTATTCATGAACCTAAATTTTCTGATTTGTTTCTCGTAGGTTCTTGATAA 1740
Db 2690 TCTCTAATTTATTCATGAACCTAAATTTTCTGATTTGTTTCTCGTAGGTTCTTGATAA 2749

Qy 1741 GAAGAGCTCAAGTTCTCTGACATTTCTTCTGGAACTATAAGGTCAATACCGGCAAGACA 1800
Db 2750 GAAGAGCTCAAGTTCTCTGACATTTCTTCTGGAACTATAAGGTCAATACCGGCAAGACA 2809
Qy 1801 TTATGATTTCAAGTCTATATAACATCCATCGTTCTTCGAGGTACAGTTCTTCTCTCTC 1860
Db 2810 TTATGATTTCAAGTCTATATAACATCCATCGTTCTTCGAGGTACAGTTCTTCTCTCTC 2869
Qy 1861 GTCCATAGTATAACATAGGGAGGCTTAATCTTCTCTCAATGATCTTTGTGTGTTGG 1920
Db 2870 GTCCATAGTATAACATAGGGAGGCTTAATCTTCTCTCAATGATCTTTGTGTGTTGG 2929
Qy 1921 ATATCTAACCCGAGTGGACATTTCTAGTATTTACATTTCAATGATCTTTGTGTGTTGG 1980
Db 2930 ATATCTAACCCGAGTGGACATTTCTAGTATTTACATTTCAATGATCTTTGTGTGTTGG 2989
Qy 1981 GTTGTGTTGTTATTTCCAAAGGTATGGGAAAAAGCTGAGGAAATTTCTGCTGAACGATTCGA 2040
Db 2990 GTTGTGTTGTTATTTCCAAAGGTATGGGAAAAAGCTGAGGAAATTTCTGCTGAACGATTCGA 3049
Qy 2041 CATAGATGGCGCAATCCCTTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC 2100
Db 3050 CATAGATGGCGCAATCCCTTAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC 3109
Qy 2101 ATCTTTTTCACACAAACTACTGAACTCAAGATTTAGTGGTTTGTATAGGGAAATTTAAAGAT 2160
Db 3110 ATCTTTTTCACACAAACTACTGAACTCAAGATTTAGTGGTTTGTATAGGGAAATTTAAAGAT 3169
Qy 2161 GATTTTCTTTTTCACAGATTTCAATCCATTCAGTGGAGGGCTTAGAAAAATGTGTAGGCG 2220
Db 3170 GATTTTCTTTTTCACAGATTTCAATCCATTCAGTGGAGGGCTTAGAAAAATGTGTAGGCG 3229
Qy 2221 ATCAGTTTGCATTTGATGAGGCAATTTGGCACTCGCGGTGTTTCTCAGCGGTTTAAACG 2280
Db 3230 ATCAGTTTGCATTTGATGAGGCAATTTGGCACTCGCGGTGTTTCTCAGCGGTTTAAACG 3289
Qy 2281 TTGAGCTGTTCTCTCATCAGACCAATTTAGCATGACCAACAGGAGCAACCATACACACCA 2340
Db 3290 TTGAGCTGTTCTCTCATCAGACCAATTTAGCATGACCAACAGGAGCAACCATACACACCA 3349
Qy 2341 ATGTATGCCAATGTTTCTCAGCTCGAGAGATTAATGAGAGTGTCTGTTTGTGTGAATG 2400
Db 3350 ATGTATGCCAATGTTTCTCAGCTCGAGAGATTAATGAGAGTGTCTGTTTGTGTGAATG 3409
Qy 2401 ATTCCAAATTTCTTAATGCTGATATTTCAATTTTCAGGATTTGTATATGAAGTGAAGCA 2460
Db 3410 ATTCCAAATTTCTTAATGCTGATATTTCAATTTTCAGGATTTGTATATGAAGTGAAGCA 3469
Qy 2461 AAGGTAA 2467
Db 3470 AAGGTAA 3476

RESULT 3

ABE16911 standard; cDNA; 2467 BP.

AC ABE16911;

DT 08-SEP-2005 (first entry)

DE Thale cress mutant LUT1 (cytochrome P450 97C1) cDNA, SEQ ID NO: 7.

XX Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C1;

XX cytochrome P450 97C1; ss; mutant.

OS Arabidopsis thaliana.

OS Synthetic.

XX US2005150002-A1.

XX 07-JUL-2005.

QY	1681	TCTCTAATTATTTCATGAACCTAAATTTTCTGATTGATTTCTTCTGGTAGGCTTGATAA	1740
DB	1681	TCTCTAATTATTTCATGAACCTAAATTTTCTGATTGATTTCTTCTGGTAGGCTTGATAA	1740
QY	1741	GAAGAGCTCAAGTTCCTGACATTTCTCTGGGAACTATTAAGGTCAAATACCGGACAAAGACA	1800
DB	1741	GAAGAGCTCAAGTTCCTGACATTTCTCTGGGAACTATTAAGGTCAAATACCGGACAAAGACA	1800
QY	1801	TTATGATTTTCAGTCTATTAACATCCATCGTTCTTCCGAGGTACAGTTCTCTTCTCTCTCTC	1860
DB	1801	TTATGATTTTCAGTCTATTAACATCCATCGTTCTTCCGAGGTACAGTTCTCTTCTCTCTCTC	1860
QY	1861	GTCCATAGTATATAACATAGGGGAGCCTAAATCCCTTCTCTTCAATGATCTTTGTGTGGTTCGG	1920
DB	1861	GTCCATAGTATATAACATAGGGGAGCCTAAATCCCTTCTCTTCAATGATCTTTGTGTGGTTCGG	1920
QY	1921	ATATCTAACCGGAGTGGACATTCCTAGTATTAATCATTCATGCGCCACATTTCTTATGTGTTT	1980
DB	1921	ATATCTAACCGGAGTGGACATTCCTAGTATTAATCATTCATGCGCCACATTTCTTATGTGTTT	1980
QY	1981	GTTGTTTGTATTTCCTAAAGGTATGGGAAAAGCTGAGGAAATTTCTGCTGAACGATTCGA	2040
DB	1981	GTTGTTTGTATTTCCTAAAGGTATGGGAAAAGCTGAGGAAATTTCTGCTGAACGATTCGA	2040
QY	2041	CATAGATGGCGCAATCCCTAAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC	2100
DB	2041	CATAGATGGCGCAATCCCTAAACGAAACAAACACTGATTTCAAGTAAACTCAGTAGAACAC	2100
QY	2101	ATCTTTTGGACAAAACCTACTGAATCAAGATTAGTGGTTTTGATTAGGGAATTTAAAGAT	2160
DB	2101	ATCTTTTGGACAAAACCTACTGAATCAAGATTAGTGGTTTTGATTAGGGAATTTAAAGAT	2160
QY	2161	GATTTTCTTTTTCACACAGATTATCCCATTCAGTGGAGGGCCCTAGAAAATGTAGGCG	2220
DB	2161	GATTTTCTTTTTCACACAGATTATCCCATTCAGTGGAGGGCCCTAGAAAATGTAGGCG	2220
QY	2221	ATCAGTTTGGCATTCGATGGAGCAATTCGTGGCACTCGCGGTGTTTCTTTCAGCGGTAAACG	2280
DB	2221	ATCAGTTTGGCATTCGATGGAGCAATTCGTGGCACTCGCGGTGTTTCTTTCAGCGGTAAACG	2280
QY	2281	TTGAGCTGGTTCCTGATCAGACCAATTAGCATGACCAAGGAGGCAACATACACACACCA	2340
DB	2281	TTGAGCTGGTTCCTGATCAGACCAATTAGCATGACCAAGGAGGCAACATACACACACCA	2340
QY	2341	ATGTTATGCCAATGTTCTCAGACTCGAGAGATTATGAGAGTGTCTGTTTGTTTAGATG	2400
DB	2341	ATGTTATGCCAATGTTCTCAGACTCGAGAGATTATGAGAGTGTCTGTTTGTTTAGATG	2400
QY	2401	ATTTCCAAATTTCTTAATGCTGATATTTTCAATTTTCAGGGAATGTATATGAAGGTGAGCCA	2460
DB	2401	ATTTCCAAATTTCTTAATGCTGATATTTTCAATTTTCAGGGAATGTATATGAAGGTGAGCCA	2460
QY	2461	AAGGTAA 2467	
DB	2461	AAGGTAA 2467	

RESULT 4

ADT15447

ID ADT15447 standard; cDNA; 3377 BP.

PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAJ/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
FI WPI; 2004-180133/17.
DR
XX
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX
PS Claim 1; SEQ ID NO 32608; 15pp; English.
XX
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
XX invention.
XX
SQ Sequence 1804 BP; 464 A; 392 C; 472 G; 476 T; 0 U; 0 Other;

Query Match 7.9%; Score 195.8; DB 13; Length 1804;
Best Local Similarity 63.2%; Pred. No. 1.6e-42;
Matches 371; Conservative 0; Mismatches 127; Indels 89; Gaps 1;

QY 647 GCGAGGCGTAGAGCGGTGTTCCATCGCTTCACAGGAGGTATTTGCTGTGATTGTGGA 706
DB 394 GGTGAGACGACAGCTGTTGTGCTTCCCTGCACAAAGATTTCGTCCATTATCGTGA 453

QY 707 GAGAGTATTCGCAATGTCAGAGAGGCTTGTGAGAAAGTTGCAGCCTTATGCAGAAGA 766
DB 454 CRAAGTATTTGTAATGCTGAGAGACTATAGACAACACTGAGCCATATGCTTTGAG 513

QY 767 CGGAAGTGTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTGTGATTCATGGGTT 826
DB 514 TGGGGAACTGTCAATATGGAAGCGAGGTTTCTCAGTTGACATGGATGATGATGGTTT 573

QY 827 GTCTCTTTTAACTACAAATTCGATCTTTGACTACTGATGATCCCTGATTAAGCTGT 886
DB 574 ATCATTTGTTCACTACAAATTTGATTCCTTCACACAGATAGTCCCTGATGATGCTGT 633

QY 887 TTACACTGCTCTTAAAGAGCTGAGCTTCGTTCTTACTGATCTTCTGCCATATTTGGAAGGC 946
DB 634 TTATATCTGACTCAAGAGACAGAGTTCTGTTCTACAGATCTTTTGCCATCTGGAAGGT 693

QY 947 AAGTTTCTGTGTTTTTTTCTGTGGTTTGTGATTGTGTGGAACAATGGATTCTTGTAA 1006
DB 694 TGGTTTC----- 700

QY 1007 TTGAGAGGGTTGGTTGTTTTTTTTCAGATCGATGCTATTGTGTAAGATAGTCCCGAGACAG 1066

Db 701 -----TTGTCAAGATAATCCCAAGACAG 724

QY 1067 GTGAAGCTGAAGAGCTGTAACTTTTGTATAGGGAACCTGTTGAAGACCTTTATTGCTAAG 1126
Db 725 ATAAAGGACGAGAAATGCGGTTTACGATTATTAAGGAACACTGTTGAAGAGCTGATTATGAAG 784

QY 1127 TGTAAAGAAATTTGCGAAAGAGAGGCGGAAGAATCAATGATGAGGAGTATGTAATGAT 1186
Db 785 TGTAAAGAAATAGTGGAGCTGAAATGAACAGATTGAGGCTGAGGATATGTAACGAA 844

QY 1187 GCTGACCAAGTATCTCGGTTTCTGCTTCAAGCAGAGAGAGGT 1233
Db 845 GGGGATCCTAGCATTTACGCTTCTCTACTTGTAGCCGAGATGAGGT 891

RESULT 6
AEB16926
ID AEB16926 standard; cDNA; 1086 BP.
XX
AC AEB16926;
XX
DT 08-SEP-2005 (first entry)
XX
DE Wheat cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 24.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
XX
OS Triticum aestivum.
XX
FH Location/Qualifiers
FT 1..1086
FT /*tag= a
FT /product= "wheat cytochrome P450 97C (CYP97C) protein"
FT /partial
FT /note= "No stop codon"
XX
PN US2005150002-A1.
XX
PD 07-JUL-2005.
XX
PF 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
DR WPI; 2005-487984/49.
DR P-PSDB; AEB16920.
DR GENBANK; CA497665, BG906289, CA742365, CA742792.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 24; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using lutein hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is wheat cytochrome P450 monooxygenase
CC (CYP97C) cDNA.
XX
SQ Sequence 1086 BP; 243 A; 298 C; 295 G; 250 T; 0 U; 0 Other;

```
Query Match      7.6%; Score 187.2; DB 14; Length 1086;
Best Local Similarity 61.9%; Pred. No. 2.9e-40;
Matches 369; Conservative 0; Mismatches 138; Indels 89; Gaps 1;

QY 647 GCGAGGCGTAGAGCGGTGTTCCATCGCTTCACAGGAGGTATTTGCTGTGATTTGCGA 706
DB 489 GGTGAGACGTAGACGCGTGTACCATCTCTACACAAAGATTTCTCTGATTAATGTCGA 548

QY 707 GAGAGTATTTCTCAATGTCAGAGAGCGCTTGTGTGAGAAAGTTGCGAGCCTTATGCGAAGA 766
DB 549 TAAAGTGTCTGATAATGTCTGAGAGATTTGTGTGAAAGAGCTCGAGACTTATGCTTTGAG 608

QY 767 CGAAGTGTCTGATAATGTCAGAGCGAGTCTCTCGATGACACATTTGATGCTATTTGGGTT 826
DB 609 TGGTGAACCTGTAAATATGGAAGCGAGGTTTCTCAATGACATTAGATGATGATTTGGTTT 668

QY 827 GTCTCTTTTAACTCAATTTCTGATTTCTTGTACTGATGATGATGATGATGATGATGATG 886
DB 669 ATCCTTGTTCACACTTGTGATTTCCCTCATCATCAGATAGTCTGTTATGATGCTGT 728

QY 887 TTACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGGC 946
DB 729 TTACTGCTCTTAAAGAGCTGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGGC 783

QY 947 AAGTTTCTGTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1006
DB 784 ----- 783

QY 1007 TTGAGAGGTTTGGTTGTTTCTTTCAGATCGATGCAATTGTGTAGATAGTCCCGAGACG 1066
DB 784 -----CAGATCGATTTGCTGTGCAAGATTTGCTCTAGACAG 819

QY 1067 GTGAAGCTGAAAGCGTGTAACTTTGATAAGGGAACCTGTTGAAGACCTTATTTGCTAAG 1126
DB 820 ATAAAGCGGAAAGAGCAGTTTAACACATAGGAATACCGTTGAAGAGCTAATACAAA 879

QY 1127 TGTAAAGAAATTTGTCGAAGAGAGCGGAAAGAAATCAATGATGAGAGATGATTAATGAT 1186
DB 880 TCGAAGGCAATCGTAGATGCTGAAATGAAACAGATTTGAGGGTGAAGAAATATGTAATGAG 939

QY 1187 GCTGACCAAGTATCTCGTTCTTCTGCTTCAACAGAGAGAGGTTTAACTTT 1242
DB 940 GCAGATCCTAGCATCCTCGGTTTCTTACTTGTAGCCGTTGAAGAGGTGAGCAGTTT 995

RESULT 7
ID AEB16925
XX
AC AEB16925;
XX
DT 08-SEP-2005 (first entry)
XX
DE Barley cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 23.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
XX cytochrome P450 97C; gene; ss.
XX
OS Hordeum vulgare; subsp. vulgare.
XX
OS Hordeum vulgare; subsp. spontaneum.
XX
FH Key
CDS Location/Qualifiers
FT 1..1638
FT /*tag= a
FT /product= "Barley cytochrome P450 97C (CYP97C) protein"
FT /transl_except= (pos:1585..1587, aa:Xaa)
FT /note= "Xaa may be any naturally occurring amino acid"
XX
XX US2005150002-A1.
PN
XX
XX 07-JUL-2005.
PD
XX
```

PP 02-JAN-2004; 2004US-00751235.
XX
PR 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIM/) KIM J.
XX
PI Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
DR P-PSDB; AEB16919.
DR GENBANK; EMB16653, BU987393, CA023004.
DR DDBJ; AV835803.
XX
PT New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
PS Claim 10; SEQ ID NO 23; 135pp; English.
XX
CC The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using luteal epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is barley cytochrome P450
CC monooxygenase (CYP97C) cDNA.
XX
SQ Sequence 1638 BP; 406 A; 403 C; 427 G; 401 T; 0 U; 1 Other;
Query Match 7.6%; Score 187.2; DB 14; Length 1638;
Best Local Similarity 61.9%; Pred. No. 3.5e-40;
Matches 369; Conservative 0; Mismatches 138; Indels 89; Gaps 1;

QY 1187 GCTGACCCAGTATCTCGGTTCTTGTGCTGCAAGCAGAGAGGTTTAACTTT 1242
 DB 925 GCAGATCTAGCATCTCGGTTTCTTACTTGTCTAGCCGTGAAGAGGTTCAGCAGTTT 980

RESULT 8
 AEB16924
 ID AEB16924 standard; cDNA; 1686 BP.
 AC AEB16924;
 XX
 DT 08-SEP-2005 (first entry)
 XX
 DE Rice cytochrome P450 97C2 (CYP97C2) cDNA, SEQ ID NO: 22.
 XX
 KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C2;
 KW cytochrome P450 97C2; gene; ss.
 XX
 OS Oryza sativa; japonica cultivar-group.
 XX
 FH Key Location/Qualifiers
 CDS 1..1686
 FT /*tag= a
 FT /product= "Rice cytochrome P450 97C2 (CYP97C2) protein"
 XX
 US2005150002-A1.
 XX
 PD 07-JUL-2005.
 XX
 PF 02-JAN-2004; 2004US-00751235.
 XX
 PR 02-JAN-2004; 2004US-00751235.
 XX
 PA (DELL/) DELLAPENNA D.
 PA (TIAN/) TIAN L.
 PA (KIMJ/) KIM J.
 XX
 PI Dellapenna D, Tian L, Kim J;
 XX
 DR WPI; 2005-487984/49.
 DR P-PSDB; AEB16918.
 DR DBU; AK065689.
 XX
 PT New expression vector comprising a nucleic acid sequence encoding a
 PT polypeptide having monooxygenase P450 activity, useful in altering the
 PT carotenoid production in a plant for enhancing production of specific
 PT carotenoid compounds.
 XX
 PS Claim 10; SEQ ID NO 22; 135pp; English.
 XX
 CC The present invention relates to genes, proteins and methods comprising
 CC carotenoid monooxygenases in the cytochrome P450 family. The invention
 CC also relates to altering carotenoid ratios in plants and microorganisms
 CC using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
 CC invention is useful in altering the carotenoid production in a plant for
 CC enhancing production of specific carotenoid compounds that are potent
 CC antioxidants. The present sequence is rice cytochrome P450 monooxygenase
 CC (CYP97C2) cDNA.
 XX
 SQ Sequence 1686 BP; 393 A; 415 C; 459 G; 419 T; 0 U; 0 Other;

Query Match 7.4%; Score 183; DB 14; Length 1686;
 Best Local Similarity 61.8%; Pred. No. 5e-39;
 Matches 363; Conservative 0; Mismatches 135; Indels 89; Gaps 1;

QY 647 GCGAGCGGTAGAGCGGTTCATCGCTTCACAGGAGTATTCTGTGTGATTCTGGA 706
 DB 522 GGTGAGACGTCGATCAGTTGTACCATCTCTACAAACGATTCTCTCGGTGATGTTGA 581
 QY 707 GAGAGTATTCTGCAAAATGTCAGAGAGGCTTGTGTGAGAAGTTTGCAGCCTTATGCAAGA 766
 DR 582 CAGAGTTTTTGTAAATGTCTGAGAGATTAGTGGAGAGCTTGAGACATCTCTTTAAG 641

QY 767 CGGAAGTCTCTGTAATATGAAGCGAAGTCTCTCAGATGACACATGATGATGGGTT 826
 DB 642 TGGCAAACTGTAATATGAAGCAAGGTTCTCTCAATGATTTAGATGATGGTTT 701
 QY 827 GTCTCTTTTAACTACAAATTTTCAATTTTGTGACTACTGATAGTCTGTGATGAGAGCTGT 886
 DB 702 GTCTCTTTTCAATTTTCAATTTTGTATTCCTCAGATCAGATAGCCCTGTTATTTGATGCTGT 761
 QY 887 TTACACTGCTCTTAAAGAGAGCTGAGCTTCTGTTCTACTGATCTTCTGCCATATTTGGAAGC 946
 DB 762 TTACACTGCACTCAAGGAAGCAGAACTTCGTTCTACAGATCTTTTACCATCTGGA---- 817
 QY 947 AAGTTTCTGCTGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1006
 DB 818 ----- 817
 QY 1007 TTGAGAGGTTTGGTTGTTTCTTTCAGATCGATGCAATTTGTTGATGATAGTCCCGAGACAG 1066
 DB 818 -----AGATTGATTTGCTGTGCAAGATTGTTCTCTAGACAA 852
 QY 1067 GTGAAGCTGAAGAGGCTGTAACCTTTGATAAGGGAACCTGTTGAAGACCTTATTCGTAAG 1126
 DB 853 ATAAAGCAGAAAAGGCGAGTTAAACATCATCAGGAATACCGTTGAGGACCTAATTACCAA 912
 QY 1127 TGTAAAGAAATTTGTAAGAGAGAGCGGCGAAAGAAATCAATGATGAGGAGTATGTAATGAT 1186
 DB 913 TGCAAGAAGATTGTAGATGCTGAGAAATGACAAATTTGAGGGGTGAGGAATATGTAATGAG 972
 QY 1187 GCTGACCCAGTATCTCGTTCCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1233
 DB 973 GCAGACCTTAGCATCTCGGATTCCTACTTCTGCTAGCCGTGAAGAGGT 1019

RESULT 9
 AEB16929
 ID AEB16929 standard; cDNA; 624 BP.
 XX
 AC AEB16929;
 XX
 DT 08-SEP-2005 (first entry)
 XX
 DE Sunflower cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 27.
 KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
 KW cytochrome P450 97C; gene; ss.
 XX
 OS Helianthus annuus.
 XX
 FH Key Location/Qualifiers
 CDS 1..624
 FT /*tag= a
 FT /product= "Sunflower cytochrome P450 97C (CYP97C)
 FT protein"
 FT /partial
 FT /note= "No start and stop codons"
 XX
 US2005150002-A1.
 XX
 PD 07-JUL-2005.
 XX
 PF 02-JAN-2004; 2004US-00751235.
 XX
 PR 02-JAN-2004; 2004US-00751235.
 XX
 PA (DELL/) DELLAPENNA D.
 PA (TIAN/) TIAN L.
 PA (KIMJ/) KIM J.
 XX
 PI Dellapenna D, Tian L, Kim J;
 XX
 DR WPI; 2005-487984/49.
 DR P-PSDB; AEB16923.
 DR GENBANK; BQ971938.

```
XX New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
PS Claim 10; SEQ ID NO 27; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using lutein-epoxide-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is sunflower cytochrome P450
CC monooxygenase (CYP97C) cDNA.
XX Sequence 624 BP; 186 A; 110 C; 144 G; 184 T; 0 U; 0 Other;
Query Match 6.9%; Score 169.6; DB 14; Length 624;
Best Local Similarity 69.6%; Pred. No. 1.5e-35;
Matches 245; Conservative 0; Mismatches 104; Indels 3; Gaps 1;
QY 608 TTCTCTATTGGGCTTAGCTGAAGTCTGATTTTATCATAGGCGAGCGTAGAGCGGTGT 667
Db 117 TTCTGGTTTGCATTCCTGAAGGCTCTCTTTGGAC---TGCAGGCGGAGGCGGTGTAGT 173
QY 668 TCCATCGCTTCCACGAGGATTTTGTCTGTGATTTGGAGAGAGTATCTGCAAAATGTGC 727
Db 174 TCCATCACTTCCACAGAGTACTTATCATGATATAGTTGATCTGTATTTGCCAATGCTC 233
QY 728 AGAGAGGCTTGTGAGAGTTCGACGCTTATGCGAAGACGGAAGTGTGTAATATGGA 787
Db 234 CGAGAGGCTTGTGGAAGAGCTAAGATCATACGACGCGAGTGTGTTAATATGGA 293
QY 788 AGCGAAGTCTCTCAGATGACACTTGATGTCATTCGGTTGTCTCTTTTAACTACAATT 847
Db 294 GCAACAGTTTTCGCGAGTTAAACCTTGATGTTATCGGTCTAGCCGTATTTAACTACAATT 353
QY 848 CGATTCTTTGACTACTGATAGTCTCTGTCATTTGAAGCTGTTTACACTGTCTTAAAGAGC 907
Db 354 TGACTCACTTAGCGCGAGTAGTCTCTGTAATGTAATCTGTTATACCGCACTAAAGAGC 413
QY 908 TGAGCTTCGTTCTACTGATCTTCTGCCATATTTGGAAGGCAAGTTTCTCTGTGT 959
Db 414 TGAAGCCCGTTTCAACTGATCTTTTGCATATTTGGAAGATAGTGGGTATGT 465
RESULT 10
ABE16927
ID ABE16927 standard; cDNA; 839 BP.
XX
AC ABE16927;
XX
DT 08-SEP-2005 (first entry)
XX
DE Tomato cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 25.
XX
DE Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; 88.
XX
XX Lycopersicon esculentum.
OS
XX Key Location/Qualifiers
XX CDS 1..837
XX /*tag= a
XX /product= "Tomato cytochrome P450 97C (CYP97C) protein"
XX /partial
XX /note= "NO start and stop codons"
XX
XX US2005150002-A1.
XX
PD 07-JUL-2005.
```

```
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
PA (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI: 2005-487984/49.
XX P-PSDB; ABE16921.
XX GENBANK; BG643819.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX Claim 10; SEQ ID NO 25; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using lutein-epoxide-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is tomato cytochrome P450
CC monooxygenase (CYP97C) cDNA.
XX
XX Sequence 839 BP; 251 A; 160 C; 190 G; 238 T; 0 U; 0 Other;
Query Match 6.5%; Score 161; DB 14; Length 839;
Best Local Similarity 61.7%; Pred. No. 3.8e-33;
Matches 329; Conservative 0; Mismatches 115; Indels 89; Gaps 1;
QY 717 TGCAAATGTGAGAGAGCGCTTGTGAGAAAGTTGCGACCTTATGCGAAGACGGAAGTGT 776
Db 1 TGCAGATGTCTGAGAGAAATGTTGAGAAACCTTTTACCTGATGCAATTTCTGGCTCTGCA 60
QY 777 GTGATATGGAAGCGAAGTTCCTCAGATGACACTTGATGTCATTTGGGTGTCTCTTTT 836
Db 61 GTGATATGGAAGCGAAGTTCCTCAGATGACACTTGATGTCATTTGGGTGTCTCTTTT 120
QY 837 AACTACAATTTTCGATTTCTTTGACTACTGATAGTCTCTGTCATTTGAAGCTGTTTACACTGT 896
Db 121 AATTACAATTTTCGATTTCTTTGACTACTGATAGTCTCTGTCATTTGAAGCTGTTTACACTGT 180
QY 897 CTTAAGAAGCTGAGCTTTCGTTCTACTGATCTTCTGCGCATATTTGGAAGCAAGTTTCTCTG 956
Db 181 CTTAAGAAGCAAGCAAGTCCGTTCACTGATTTGTTGCCATATTTG----- 225
QY 957 TGTCTTTCTGCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1016
Db 226 ----- 225
QY 1017 TTGGTTGTTTTCAGATCGATGTCATTTGTGTAGATAGTCCCGAGACAGAGTGAAGCTG 1076
Db 226 -----CAGATCAAGCTTTATGTAGTTTCATCCACGACCAATTAAGGCTG 271
QY 1077 AAAAGGCTGTAACTTTTGTATAAGGAAACTGTTGAGAGCTTATTGCTAAGTGTAAAGAAA 1136
Db 272 AGAATGCAAGTGTCTTATTAATCAGACAAACAGTTTGAAGAACTTATTTCGGAAGTGCAGAGAGA 331
QY 1137 TTGTGCAAGAGAGAGCGGCAAGATCAATGATGAGGAGTATGTAATGATGCTGACCCCA 1196
Db 332 TTGTAGAACTGAGGCGGTGAGAGGATTAATGAAGATGAGTACGTGATGATGAGATCCAA 391
QY 1197 GTATCTGCGTTTCTTGTCTTCAAGAGAGAGGTTTAAACTTTTTCCTT 1249
Db 392 GATCTCTCGATTTTTCCTGCTAGCCGTGAGGAGTTTCAAGTTTACACTT 444
```

```

RESULT 11
ID AEB16928 standard; cDNA; 531 BP.
XX
AC AEB16928;
XX
DT 08-SEP-2005 (first entry)
XX
DE Maize cytochrome P450 97C (CYP97C) cDNA, SEQ ID NO: 26.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97C;
KW cytochrome P450 97C; gene; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 1..531
FT FT /*tag= a
FT FT /product= "Maize cytochrome P450 97C (CYP97C) protein"
FT FT /partial
FT FT /note= "No start and stop codons"
XX
XX US2005150002-A1.
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
XX PA (TIAN/) TIAN L.
XX PA (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16922.
XX DR GENBANK; BE552887.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
XX polypeptide having monooxygenase P450 activity, useful in altering the
XX carotenoid production in a plant for enhancing production of specific
XX carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 26; 135pp; English.
XX
XX The present invention relates to genes, proteins and methods comprising
XX carotenoid monooxygenases in the cytochrome P450 family. The invention
XX also relates to altering carotenoid ratios in plants and microorganisms
XX using LUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
XX invention is useful in altering the carotenoid production in a plant for
XX enhancing production of specific carotenoid compounds that are potent
XX antioxidants. The present sequence is maize cytochrome P450 monooxygenase
XX (CYP97C) cDNA.
XX
XX Sequence 531 BP; 158 A; 91 C; 130 G; 152 T; 0 U; 0 Other;
XX
Query Match 5.9%; Score 144.6; DB 14; Length 531;
Best Local Similarity 62.1%; Pred. No. 9.4e-29;
Matches 300; Conservative 0; Mismatches 94; Indels 89; Gaps 1;
XX
QY 751 AGCCTTATGACAGGAGGAGTGTGTAATGGAAGCGAAGTTCCTCAGATGACAC 810
DB |||||
5 AGCCATATGTTTGAAGTGGGACCTGTCAATATGGAAGCGAGTTTCTCAGTTGACAT 64
QY 811 TTGATGTCATTTGGTGTCTCTTTTAACTACAAATTCGATTTCTTCACTGATGATC 870
DB |||||
65 TGGATGTGATTTGGTTTATCATCTTGTCACTACAAATTTGATTTCCTCACAACAGATGTC 124
QY 871 CTGTCAATTGAGCTGTTTACACTGCTTTAAAGAGCTGAGCTTCGTTCTACTGATCTTC 930
DB |||||

```

```

Db 125 CTGTCAATTGATGCTGTTTATATCTCACTCAAGAAGCAGAGCTTCGTTCTACAGATCTTT 184
QY 931 TGCCATATTCGAGGCAAGTTTCTCTGTGTTTCTGTGGTTTGTGTTGTTGTTGTTGGAACA 990
Db 185 TGCCATATTCGAGGCTGGTTTC----- 207
QY 991 ATTGGATTCCTGTTTAAATTGAGAGGGTTTGGTTGTTTTTTTTCAGATCGATGCAATGTTGTA 1050
Db 208 -----TTGTGCA 215
QY 1051 GATAGTCCGAGACAGGTGAAAGCTGAAAGGCTGTAACTTTTGAAGGAAACCTGTGA 1110
Db 216 GATAATCCCAAGACAGATATAAAGCAGAGATGCGGTTACGATTATAAGGAACACTGTGA 275
QY 1111 AGACCTTATTCCTAAGTGTAAAGAAATTGTGAAAGAGAGCGCAAGAAATCAATGATGA 1170
Db 276 AGAGCTGATTATGAGTGTAAAGAAATAGTGAAGCTGAAATGAACAGATTGAGGTTGA 335
QY 1171 GGAGTATGTAATGATGCTGACCAAGTATCTCGTTTCTTGTGTTTGAAGCAGAGAGA 1230
Db 336 GGAATATGTAACGAGGAGATCCTAGCAATCTAGCTTCTAGCTTCTAGCTAGCGAGATGA 395
QY 1231 GGT 1233
Db 396 GGT 398
XX
RESULT 12
ID AEB16943 standard; DNA; 1788 BP.
XX
AC AEB16943;
XX
DT 08-SEP-2005 (first entry)
XX
DE Thale cress cytochrome P450 97A (CYP97A) DNA, SEQ ID NO: 41.
XX
KW Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
KW cytochrome P450 97A; gene; ds.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 1..1788
FT FT /*tag= a
FT FT /product= "Thale cress cytochrome P450 97A (CYP97A)
FT FT protein"
XX
XX US2005150002-A1.
XX
XX 07-JUL-2005.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX 02-JAN-2004; 2004US-00751235.
XX
XX (DELL/) DELLAPENNA D.
XX PA (TIAN/) TIAN L.
XX PA (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
XX P-PSDB; AEB16935.
XX DR GENBANK; AY058173.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
XX polypeptide having monooxygenase P450 activity, useful in altering the
XX carotenoid production in a plant for enhancing production of specific
XX carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 41; 135pp; English.
XX

```

CC The present invention relates to genes, proteins and methods comprising
 CC carotenoid monooxygenases in the cytochrome P450 family. The invention
 CC also relates to altering carotenoid ratios in plants and microorganisms
 CC using luti epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
 CC invention is useful in altering the carotenoid production in a plant for
 CC enhancing production of specific carotenoid compounds that are potent
 CC antioxidants. The present sequence is thale cress cytochrome P450
 CC monooxygenase (CYP97A) DNA.

XX Sequence 1788 BP; 517 A; 344 C; 407 G; 520 T; 0 U; 0 Other;

Query Match 3.7%; Score 91; DB 14; Length 1788;
 Best Local Similarity 60.9%; Pred. No. 8.1e-14;
 Matches 148; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 701 TGTGGAGAGAGTATCTGCAATGTGCAGAGAGGCTTGTGAGAGTTGCAGCCTTATGC 760
 DB 645 TATGATTAGTTTATTCGGAGAGCTTCAGATAGGCTTTGTGAGAGTTGATGCTGCTGC 704

QY 761 AGAAGCGGAGTGTGTAATATGGAAGCGAAGTTCTCTCAGATGACACTTGATGTCAT 820
 DB 705 ATTGAAGGGGAGAGTAGATGGAATCACTTCTCTCGTTTGACACTTTGATATTAT 764

QY 821 TGGTGTGCTCTTTTAACTACAATTCGATTTCTGACTTCTGACTAGTCTCTGCTCAT 880
 DB 765 TGGCAAGGGGTTTCAATTAGCACTTTGACTCCCTTACTAATGATACCGGTGATCGA 824

QY 881 AGCTGTTTACACTGCTCTTAAAGAGAGCTGAGCTTCTGTTCTACTGATCTTCTGCCAT 940
 DB 825 GCGAGTGATACCTGTTCTTAAGAGAGAGCTGAAGACAGAGTGTTCACCTATTCCTGTTG 884

QY 941 GAA 943
 DB 885 GGA 887

RESULT 13
 ID AEB16942 standard; DNA; 2057 BP.
 XX AC AEB16942;
 DT 08-SEP-2005 (first entry)
 XX Thale cress cytochrome P450 97A (CYP97A) DNA, SEQ ID NO: 40.
 DE Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
 KW cytochrome P450 97A; gene; ds.
 XX Arabidopsis thaliana.

OS Key Location/Qualifiers
 FH CDS 34..1821
 FT /*tag= a
 FT /product= "Thale cress cytochrome P450 97A (CYP97A)
 FT protein"
 XX US2005150002-A1.
 XX 07-JUL-2005.
 XX 02-JAN-2004; 2004US-00751235.
 XX 02-JAN-2004; 2004US-00751235.
 XX (DELL/) DELLAPENNA D.
 XX (TIAN/) TIAN L.
 XX (KIMJ/) KIM J.
 XX Dellapenna D, Tian L, Kim J;
 XX WPI; 2005-487984/49.
 DR P-PSDB; AEB16935.

DR GENBANK; AY056446.
 XX New expression vector comprising a nucleic acid sequence encoding a
 PT polypeptide having monooxygenase P450 activity, useful in altering the
 PT carotenoid production in a plant for enhancing production of specific
 PT carotenoid compounds.

XX Claim 10; SEQ ID NO 40; 135pp; English.

XX The present invention relates to genes, proteins and methods comprising
 CC carotenoid monooxygenases in the cytochrome P450 family. The invention
 CC also relates to altering carotenoid ratios in plants and microorganisms
 CC using luti epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
 CC invention is useful in altering the carotenoid production in a plant for
 CC enhancing production of specific carotenoid compounds that are potent
 CC antioxidants. The present sequence is thale cress cytochrome P450
 CC monooxygenase (CYP97A) DNA.

XX Sequence 2057 BP; 615 A; 377 C; 457 G; 608 T; 0 U; 0 Other;

Query Match 3.7%; Score 91; DB 14; Length 2057;
 Best Local Similarity 60.9%; Pred. No. 8.7e-14;
 Matches 148; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 701 TGTGGAGAGAGTATCTGCAATGTGCAGAGAGGCTTGTGAGAGTTGCAGCCTTATGC 760
 DB 678 TATGATTAGTTTATTCGGAGAGCTTCAGATAGGCTTTGTGAGAGTTGATGCTGCTGC 737

QY 761 AGAAGCGGAGTGTGTAATATGGAAGCGAAGTTCTCTCAGATGACACTTGATGTCAT 820
 DB 738 ATTGAAGGGGAGAGTAGATGGAATCACTTCTCTCGTTTGACACTTTGATATTAT 797

QY 821 TGGGTTGTCTCTTTTAACTACAATTCGATTTCTTGTGACTACTGATAGTCTGTCATTGA 880
 DB 798 TGGCAAGGGGTTTCAATTAGCACTTTGACTCCCTTACTAATGATACCGGTGATCGA 857

QY 881 AGCTGTTTACACTGCTCTTAAAGAGAGCTGAGCTTCTGTTCTACTGATCTTCTGCCAT 940
 DB 858 GCGAGTGATACCTGTTCTTAAGAGAGAGCTGAAGACAGAGTGTTCACCTATTCCTGTTG 917

QY 941 GAA 943
 DB 918 GGA 920

RESULT 14
 ID AEB16947 standard; cDNA; 1278 BP.
 XX AC AEB16947;
 DT 08-SEP-2005 (first entry)
 XX Soybean cytochrome P450 97A (CYP97A) cDNA, SEQ ID NO: 45.
 DE Pigment; metabolic engineering; antioxidant; transgenic plant; CYP97A;
 KW cytochrome P450 97A; gene; ss.
 XX Glycine max.

OS Key Location/Qualifiers
 FH CDS 1..1278
 FT /*tag= a
 FT /product= "Soybean cytochrome P450 97A (CYP97A) protein"
 FT /partial
 FT /note= "No start codon"
 XX US2005150002-A1.
 XX 07-JUL-2005.
 XX 02-JAN-2004; 2004US-00751235.
 XX P-PSDB; AEB16935.

PR 02-JAN-2004; 2004US-00751235.
XX (DELL/) DELLAPENNA D.
PA (TIAN/) TIAN L.
PA (KIMJ/) KIM J.
XX
XX Dellapenna D, Tian L, Kim J;
XX
XX WPI; 2005-487984/49.
DR P-PSDB; ABB16938.
DR GENBANK; BF425906, BF596805, AW704660, AW704625, BI470164, BQ296458,
DR BMS92469, AI938600, AI938382, BU544173, BI471346, CD410775, BF598710,
DR BG154747.
XX
XX New expression vector comprising a nucleic acid sequence encoding a
PT polypeptide having monooxygenase P450 activity, useful in altering the
PT carotenoid production in a plant for enhancing production of specific
PT carotenoid compounds.
XX
XX Claim 10; SEQ ID NO 45; 135pp; English.
PS
XX The present invention relates to genes, proteins and methods comprising
CC carotenoid monooxygenases in the cytochrome P450 family. The invention
CC also relates to altering carotenoid ratios in plants and microorganisms
CC using IUT1 epsilon-hydroxylases and/or CYP97A beta-hydroxylases. The
CC invention is useful in altering the carotenoid production in a plant for
CC enhancing production of specific carotenoid compounds that are potent
CC antioxidants. The present sequence is soybean cytochrome P450
CC monooxygenase (CYP97A) cDNA.
XX
XX Sequence 1278 BP; 389 A; 247 C; 297 G; 345 T; 0 U; 0 Other;
SQ
Query Match 3.7%; Score 90.8; DB 14; Length 1278;
Best Local Similarity 59.2%; Pred. No. 7.8e-14;
Matches 174; Conservative 0; Mismatches 117; Indels 3; Gaps 1;
QY 651 AGCGCTAGAGCGGTGTCATCGCTTACAGGAGGTATTTGTCGTGATTTGGAGAGA 710
Db 43 AGACGTCGTGTATAGTCCAGCATTTGCCACGAGATGATGTAGCAGCTATGATTG---GC 99
QY 711 GTATTCTGCAATGTCGACGAGGCTGTTGAGAGTTGTCAGCCCTTATGCGAAGACGGA 770
Db 100 CTTTTCGGACAAGCTGCAGATAGGCTCTCCGAGAGCTAGATGCTGTCATCCGATGGA 159
QY 771 AGTGCTGTCAATATGGAAGCGAAGTTCTCTCAGATGACACTTGATGTCATTTGGTGTCT 830
Db 160 GAAGATGTTGAGATGGAATCACTTTTCTCTCGATTGACCTTGGACATCATTTGCCAAGCA 219
QY 831 CTTTTTAACTACAAATTCGATTTCTTGGACTACTGATGCTCTGTCATTTGAAGCTGTTTAC 890
Db 220 GTATTCAATATGATTTTGATAGTTTATCAATATGACACTGGTATAGTTGAGGCTGTTTAT 279
QY 891 ACTGCTCTTAAGAGAGCTGAGCTTCCTTCTACTGATCTTCTGCCATATTGGAAG 944
Db 280 ACTGTACTGAGAGAAGCAGAGATCGAAGTGTGCTCCAAATCCAGTCTGGGAG 333
RESULT 15
AD082389 standard; cDNA; 1078 BP.
XX
XX AD082389;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polynucleotide seqid 1109.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;

yield; plant growth; plant development; seed oil; protein yield;
protein content; gene; ss.
Unidentified.
US2004034888-A1.
19-FEB-2004.
28-APR-2003; 2003US-00425114.
06-MAY-1999; 99US-00304517.
05-NOV-2001; 2001US-00985678.
(LIUJ/) LIU J.
(ZHOU/) ZHOU Y.
(KOVA/) KOVALIC D K.
(SCRE/) SCREEN S E.
(TABA/) TABASKA J E.
(CAOY/) CAO Y.
Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
WPI; 2004-180133/17.
New recombinant DNA construct, useful for improving plant tolerance to
cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
pests, for conferring increased resistance to plant disease, or for
improving yield.
Claim 1; SEQ ID NO 1109; 15pp; English.
The invention describes a recombinant DNA construct comprising a
polynucleotide consisting of a sequence encoding an amino acid sequence
available in electronic form from the US patent office at
ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
of the invention are also useful in physical arrays of molecules and as
plant breeding markers. The recombinant DNA construct is useful for
improving plant tolerance to cold, heat, drought, herbicides, extreme
osmotic conditions, pathogens or pests, for manipulating growth rate in
plant cells by modification of the cell cycle pathway, for conferring
increased resistance to plant disease, for producing galactomannan,
lignin or plant growth regulators, for increasing the rate of homologous
recombination in plants, for improving yield by modification of
photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
or by providing improved plant growth and development under at least one
stress condition or for modifying seed oil or protein yield and/or
content. This sequence represents a plant full length insert
polynucleotide that can be used in the recombinant DNA construct of the
invention.
Sequence 1078 BP; 332 A; 180 C; 253 G; 313 T; 0 U; 0 Other;
Query Match 3.5%; Score 87.4; DB 13; Length 1078;
Best Local Similarity 69.8%; Pred. No. 6.1e-13;
Matches 118; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 2177 CAGATTTCATCCATTTCAGTGGAGGCGCTAGAAATGTGTAGCGATCAGTTTGCAATTGAT 2236
Db 186 CAGGTTTATTCATTTCAGTGGGGGCTCTCGGAAATGTCTGGAGATCAATTGCTCTCCT 245
QY 2237 GGAGGCAATTTGGCACTCGCGTGTCTTTCAGCGGTTTAAACAGTTGAGCTGGTTCCTGA 2296
Db 246 AGAAGCAATAGTGGCACTTGCAGTTGTGTTCAGAAGATGATATCAGCTTGTGCCAGA 305
QY 2297 TCAGACCAATTAGCATGACCAAGGAGCAACCATACACACCACCAATGTA 2345
Db 306 TCAAAAGATTAAACATGACTACTGAGCTACAAATTCATACACACAGTGA 354
Search completed: January 13, 2006, 01:04:32
Job time : 1451 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 23:46:35 ; Search time 9655 Seconds
(without alignments)
11954.814 Million cell updates/sec

Title: US-10-751-235-5
Perfect score: 2467
Sequence: 1 atggagttcttactcttttc.....tgaaggtgagccaaaggttaa 2467

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_ges1:*
10: gb_ges2:*
11: gb_ges3:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	791	32.1	1709	4	CNS0A665
2	723.2	29.3	1664	4	CNS0A5YJ
3	701.4	28.4	736	10	CW839414
4	499	20.2	616	1	AV827314
5	473	19.2	977	10	CL469740
6	442	17.9	657	9	CB255014
7	429.4	17.4	708	9	BH931006
8	386.6	15.7	570	9	BH584135
9	323	13.1	559	6	CB282034
10	323	13.1	716	6	CX189683
11	322.8	13.1	664	6	CD828321
12	312.4	12.7	627	6	CD828299
13	301.2	12.2	310	10	BX531442
14	228.6	9.3	858	7	CO072553
15	220.4	8.9	704	10	CZ710414
16	218.4	8.9	247	10	CW839413
17	211.8	8.6	897	8	CV886111
18	211	8.6	914	7	CK290930
19	209.6	8.5	563	2	BF096935
20	204.8	8.3	886	8	DR923995
21	203.2	8.2	604	8	CX186676
22	203	8.2	386	3	BP827113

23	202.2	8.2	642	8	CX176509
24	199.8	8.1	668	8	DR952300
25	194.2	7.9	447	1	AV422798
26	192.6	7.8	721	7	CV529617
27	191.8	7.8	469	8	CX541939
28	191.6	7.8	650	10	CZ799589
29	191.6	7.8	652	2	BI263819
30	187.2	7.6	600	5	BU987393
31	187	7.6	693	10	CW166683
32	185.6	7.5	722	8	DR403827
33	185	7.5	444	1	AI736606
34	185	7.5	511	2	BF597405
35	185	7.5	530	3	BI701721
36	182.4	7.4	589	7	CNI82757
37	181.6	7.4	440	5	BQ585965
38	180.8	7.3	480	1	AJ486563
39	180.8	7.3	596	1	AJ432371
40	179.2	7.3	733	7	CV712131
41	178.4	7.2	451	8	CX023647
42	175	7.1	689	5	BQ862275
43	173.8	7.0	941	4	CNS0A53V
44	170.4	6.9	457	1	AW617348
45	169.6	6.9	626	5	BQ971938

ALIGNMENTS

RESULT 1	CNS0A665	1709 bp	mRNA	linear	HTC 06-FEB-2004
LOCUS	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTFF64ZA07 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).				
ACCESSION	BX822785				
VERSION	BX822785.1	GI:42464524			
KEYWORDS	HTC; GSLT_CDNA.				
SOURCE	Arabidopsis thaliana				
ORGANISM	Arabidopsis thaliana (thale cress)				
REFERENCE					
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1709)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_RF/Full-length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.				
FEATURES	Location/Qualifiers				
source	1..1709				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				

CX176509	BI2_69-58
DR952300	EST114383
AV422798	AV422798
CV529617	ALV_011A
CX541939	sl3DNF0BH
CZ799589	OC_Ba017
BI263819	NF114G06P
BU987393	HF14K04r
CW166683	104_575_1
DR403827	CSAH-PNPI
AI736606	eb31b06.y
BF597405	eu97408.y
BI701721	sa119d02.
CNI82757	UCRCS04_0
BQ585965	E012531-0
AJ486563	AJ486563
AJ432371	AJ432371
CV712131	UCRP701_0
CX023647	Md1v4-404
BQ862275	QGC20K13.
CNS0A53V	Arabidops
AW617348	EST323759
BQ971938	QHB9C20.y

RESULT 2	CNS0A5YJ	1664 bp	linear	HTC 06-FEB-2004
CNS0A5YJ	Arabidopsis thaliana Full-length cDNA Complete sequence from clone			
LOCUS	GLT1PB252E04 of Flowers and buds of strain col-0 of Arabidopsis			
DEFINITION	thaliana (Thale cress).			
ACCESSION	EX8223288			
VERSION	EX8223288.1	GI:424564384		


```

Db      297  AAAGCTGAGGAATTTCTCCCTGAAGATTCGACATAGATGGCGCAATCCCTAACGAACA 238
Qy      2069  AACACTGATTTCAAGTAAACTCAGTAGAACACATCTTTTGTACACAAACTACTGAATCAAG 2128
Db      237  AACACTGATTTCAAGTAAACTCAGTAGAACACATCTTTTGTACACAAACTACTGAATCAAG 178
Qy      2129  ATTAGTGCTTTGATTAGGGAATTTAAAGATGATTTCTTTTTCACAGATTCATCCC 2188
Db      177  ATTAGTGCTTTGATTAGGGAATTTAAAGAGATTTTCTTTTTCACAGATTCATCCC 118
Qy      2189  ATTCACTGGAGGCGCTAGAAA 2210
Db      117  ATTCACTGGAGGCGCTAGAAA 96

RESULT 6
LOCUS   CB255014
DEFINITION 54-E018363-019-007-I14-T7R MP1Z-ADIS-019 Arabidopsis thaliana cDNA
clone MP1Zp768L147Q 5-PRIME, mRNA sequence.
ACCESSION CB255014
VERSION   CB255014.1 GI:56915639
KEYWORDS EST.
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 657)
REFERENCE
AUTHORS Jakoby,M., Stracke,R., Soerensen,T.R. and Weisshaar,B.
TITLE Arabidopsis thaliana cDNA library enriched in transcription factors
JOURNAL Unpublished (2003)
COMMENT Contact: Weisshaar B
ADIS DNA core facility at MP1Z
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaar@piz-koeln.mpg.de
Insert Length: 657 Std Error: 0.00
Plate: 7 row: L column: 14
Seq primer: T7R; CTAATACGACTCACTATAGGA.

FEATURES
source
1..657
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="At7"
/db_xref="GABI:597448"
/db_xref="taxon:3702"
/clone="MP1Zp768L147Q"
/tissue_type="hypocotyl"
/dev_stage="tissue culture"
/lab_host="E. coli DH5alpha mcr"
/clone_lib="MP1Z-ADIS-019"
/notes="Vector: pSPORT1; Site 1: NotI primer adapter;
Site 2: SalI primer adapter; RNA from cell culture (At7) 5
days after inoculation treated with 0.002 mM cycloheximid
for 2 h in the dark. Sequence submission managed by
RZPD/GABI-Primary database: http://gabi.rzpd.de. This
clone is available from RZPD; contact RZPD (clone@rzpd.de)
for further information."

ORIGIN
Query Match 17.9%; Score 442; DB 6; Length 657;
Best Local Similarity 79.5%; Pred. No. 4.8e-107;
Matches 630; Conservative 0; Mismatches 10; Indels 152; Gaps 3;

Qy      155  AATCCCAATCATGGGTGAGTCCGATGGCTCAACACTCACTCTGACCTTCTCTCAG 214
Db      5  AATCCCAATCATGGGTGAGTCCGATGGCTCAC-ACACTCACTCTGACCTTCTCTTTT 63
Qy      215  GAAAAACGACGATCAGGTATACCAATCGCAACGAGCTCGACGATGTCGCTGATC 274
Db      64  GACAAACGACGATCAGGTATACCAATCGCAACGAGCTCGACGATGTCGCTGATC 123

```

```

Qy      275  TCCTCGAGGTGCTCTCTCTTACCTCTCTACAAATGGATGAATGAGTACGAGCCCATTT 334
Db      124  TCCTCGAGGTGCTCTCTCTTACCTCTCTACAAATGGATGAATGAGTACGAGCCCATTT 183
Qy      335  ACCGTCCTCGCTCGTCTCGTAAATTTGTAATTTGTAAGCGACCCAGCATAGCTAAAC 394
Db      184  ACCGTCCTCGCTCGTCTCGTAAATTTGTAATTTGTAAGCGACCCAGCATAGCTAAAC 243
Qy      395  ATGTTTGTAGGAATATCCAAAGTACGCTAAAGCGTTAGTCGCTGAAAGTCTCTGAATTC 454
Db      244  ATGTTTGTAGGAATATCCAAAGTACGCTAAAGCGTTAGTCGCTGAAAGTCTCTGAATTC 303
Qy      455  TATTTGGTTTCGGGTTTCGCTATCGCTGAAGACCTCTTTGGACAGTAAATTTTCATCCTC 514
Db      304  TATTTGGTTTCGGGTTTCGCTATCGCTGAAGACCTCTTTGGACA----- 347
Qy      515  CTATCTCAATTTTGAAGTTTTTGGAAATTTGTGGAAGTAAATGTGTGACTCTCTTGTATGATA 574
Db      348  ----- 347
Qy      575  AGTAACTCTAAATTTTAGGGTTTAGATTCCAATCTCTCTATTTGGGCTTAGCTGAAGTCTG 634
Db      348  ----- 347
Qy      635  ATTTTTCATAGAGCGGCGGTAGAGCGGTGCTTCCATCGCTTTCACAGGAGGTATTTGTC 694
Db      348  -----GGAGCGGTAGAGCGGTGCTTCCATCGCTTTCACAGGAGGTATTTGTC 394
Qy      695  TGTGATTTGTGAGAGAGTATTCTGCAATGTGCAGAGAGGCTTGTGTGAGAAAGTTGAGGCC 754
Db      395  TGTGATTTGTGAGAGAGTATTCTGCAATGTGCAGAGAGGCTTGTGTGAGAAAGTTGAGGCC 454
Qy      755  TTATCGAAGACCGGAAGTGTGTGAATATGGAAGCGAAGTTCTCTCAGATCAGACTTGA 814
Db      455  TTATCGAAGACCGGAAGTGTGTGAATATGGAAGCGAAGTTCTCTCAGATCAGACTTGA 514
Qy      815  TGTCAATCGGTTGCTCTCTTTTAACTCAATTTTCCATCTTTTGAATCTACTGATGAGTCTCT 874
Db      515  TGTCAATCGGTTGCTCTCTTTTAACTCAATTTTCCATCTTTTGAATCTACTGATGAGTCTCT 574
Qy      875  CATTTGAAGCTGTTTACACTGCTCTTTAAAGAA--GCTGAGCTTCGTTCTACTGATCTCTG 932
Db      575  CATTTGAAGCTGTTTACACTGCTCTTTAAAGAAAGCTTGAGCTTCGTTCTACTGATCTCTG 634
Qy      933  CCATATTGGAAG 944
Db      635  CCATATTGGAAG 646

RESULT 7
BH931006/c
LOCUS BH931006/c
DEFINITION odii10g10.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION BH931006
VERSION BH931006.1 GI:23411072
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 708)
REFERENCE Delaunay,K., Fellw,G., Fulton,L., McCombie,W.R., Miner,T.,
AUTHORS Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE Whole genome shotgun reads from Brassica oleracea
JOURNAL Unpublished (2002)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: odii10 row: g column: 10

```

```

Seq primer: -28RppOT reverse
Class: shotgun
High quality sequence start: 81
High quality sequence stop: 434.
Location/Qualifiers
1. .708
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea TOL000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

```

ORIGIN	Query Match Best Local Similarity Matches	17.4% 81.9% 525; Conservative	Score 429.4; Pred. No. 1.2e-103; 0; Mismatches	DB 9; Length 708; Indels 15;	Gaps 2;
Qy	626	TGAAGTCTGATTTTTCATAGGCGAGCGGTAGAGCGGTGTTCCATGCTTCACAGAG	685		
Db	672	TGAATTCCTGTGTGATCTTAGGCGAGGAGAGCGGTGTTCCGTCACTTCACAAGAG	613		
Qy	686	GTATTTCTCTGTGATCTGGAGAGATTTCTGCAAAATGTGCAGAGGCTTGTGTGAA	745		
Db	612	GTACTTCTGTGATGTGGAGAGAGTGTCTGCAGATGTGCAGAGGCTCGTTGAA	553		
Qy	746	GTTCAGCCTTATGCAGAGACGGAAGTCTGTCAATATAGAACGCAAGTTCTCTCAGAT	805		
Db	552	GCTGCACCTTATGCAGTGGAGGAAAGCTGTGNAATGGAAGAGAGTTCTCTCAGTT	493		
Qy	806	GACACTTGATGTCAATGGGTGTCTCTTTTAACTACAATTCGATCTTTTGACTACTGA	865		
Db	492	AACTCTCGAGGTGATGGCTTATCTCTTTTAACTGCAACTTTGATTCACCTACCGA	433		
Qy	866	TAGTCCTGTCAATGCAAGCTGTTACAGCTCTTAAAGAGCTGAGCTCGTTCTACTGA	925		
Db	432	TAGTCCTGTCAATGCAAGCTGTTACAGCTCTCAAGAGAGCTGAGCTCGTTCTACTGA	373		
Qy	926	TCCTTCGCCATATGGAAGGCAAGTTCCTGTGTGTTTTCTGTGGTTGTGTGATGTGTG	985		
Db	372	TATCTACCCCTATGGAAGGCAAGTTGTGTGTGTTTCTGTGTGG-----C 327			
Qy	986	GAACAAATTGATCTCTGTTAATTGAGAGGGT--TTGGTTGTTTTTTTCAGATCGATGCATT	1044		
Db	326	GCAGACTTGATCTTGTGTAGCCTTTGGTTGTTTTTTTCTTTGTTTCAGATCGATGCGTT	267		
Qy	1045	GTGTAAAGATAGTCCCGAGACAGGTGAAAGCTGAAAGCGGTAACTTTGATTAAGGGAAC	1104		
Db	266	GTGTAAAGATAGTCCCGAGACAGGTGAAAGCAGAAAGGCTGTGACTTTGATTAAGGGAAC	207		
Qy	1105	TGTTGAAGACCTTATTCCTAAGTGTAAAGAAATTTGTCMAAGACGCGCAAGAAATCAA	1164		
Db	206	AGTTGAGGATCTCATTCGAAATGCAAGAGATTTGTGAAAGAGAGCGCAAGAAATCAA	147		
Qy	1165	TGATGAGGAGTATGTAAATGATGTCAGCCCAAGTATCTCGGTTTCTTGCTTGAACGAG	1224		
Db	146	TGATGAAGAGTATGTGAATGAGTCTGACCCCAAGTATCTTGGGTTCTTGCTTGAACGAG	87		
Qy	1225	AGAAGAGGTTTAACTTTTTTCTTAAAGTTTATTAAGCAAT 1265			
Db	86	AGAAGAGGTTTAACTTTCTTTTCTTTTATATTCATCCATAT 46			

RESULT 8	570 bp	DNA	linear	GSS 15-DEC-2001
BH584135/c				
LOCUS	BH584135			
DEFINITION	BOHAD90TF BOHA	Brassica oleracea	clone BOHAD90,	genomic survey sequence.

ACCESSION	BH584135
VERSION	BH584135.1
KEYWORDS	GI:17836592
SOURCE	GSS.
ORGANISM	Brassica oleracea
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
AUTHORS	1 (bases 1 to 570) Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S., Uterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
TITLE	Whole genome shotgun sequencing of Brassica oleracea and its application to gene discovery and annotation in Arabidopsis
JOURNAL	Genome Res. 15 (4), 487-495 (2005)
PUBMED	15805490
COMMENT	Other GSSs: BOHAD90TR Contact: Chris Town
FEATURES	TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org DNA is from a doubled haploid provided by Tom Osborn. Seq primer: TP Class: sheared ends. Location/Qualifiers 1..570 /organism="Brassica oleracea" /mol_type="genomic DNA" /strain="TO1000DH3" /db_xref="taxon:3712" /clone_lib="BOHAD90" /notes="Vector: pHSO1; Site 1: BatXI; 2-3 kb sheared genomic DNA inserted into pHSO1 using BatXI linkers"
ORIGIN	Query Match 15.7%; Score 386.6; DB 9; Length 570; Best Local Similarity 82.0%; Pred. No. 3.6e-92; Matches 475; Conservative 0; Mismatches 89; Indels 15; Gaps 2; QY 624 GCTGAAGTCTGATTTTACATAGCGCAGCGGTAGAGCGGTGGTTCCATCGCTTCACAGG 683 DB 565 GTTTGATTCTTGTTGATGTTAGGCAGAGAGAGCGGTGGTTCCATCTCACAG 506 QY 684 AGCATTTGTCGTGATGTGGAGAGAGTATCTGCAAATGTCAGAGAGGCTGTGTAG 743 DB 505 AGGTACTTGTCTGTGATTTGGAGAGAGTGTCTGCAGATGTCAGAGAGGCTCGTTGAG 446 QY 744 AAGTTCAGCCTTATGCAGAGACGGAAGTCTGTGAATATGMAAGCAAGTCTCTCAG 803 DB 445 AAGCTGCAGCCTTATGCAGTGGAGGAAAAGCTGTGAA CATGGAAGAAGAAGTCTCTCAG 386 QY 804 ATGACACTTGATGTCAATGGGTTGTCTTTTTTAACAATAATTCGATTTCTTGACTACT 863 DB 385 TTAACCTCTCGAGCTGATGGCTTAATCTCTTTTAACTACAACTTTGATTCACCTACC 326 QY 864 GATAGTCCTGTCAATGAAGCTGTTFACACATGCTCTTAAAGAAGCTGAGCTTGCTTCTACT 923 DB 325 GATAGTCCTGTCAATGAAGCTGTTFACACATGCTCTCAAAGAAGCTGAGCTTGCTTCTACT 266 QY 924 GATCTTCTGCCATATGGGAAGCAAGTTTCTGTGTGTTTTTCTGTGTTTGTGTGTTGTG 983 DB 265 GATATTCTACCTATTGGGAAGCAAGTTTGTGTGTTTCTTGTGTGTC----- 218 QY 984 TGGAACAATTTGATTTCTGTTAATTGAGAGGGT--TTGGTTGTTTTTTTTTCAGATCGATGCA 1042 DB 217 --GCAGACTTGGATTCCTGTTGAGCCCTTGGTTGTTTTTCTTTGTTTCAGATCGATGCG 160 QY 1043 TTGTGTAAGATAGTCCCGAGACAGGTGAAAGCTGAAAAAGCGCTGTAACTTTGTAAGGGAA 1102 DB 159 TTGTGTAAGATAGTTCGAGACAAGTGAAGCAGAAAAAGCTGTGACTTTGTAAGGGAA 100

AUTHORS Kim H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and Wing, R. A.

TITLE Global assembly of Cotton ESTs

JOURNAL Unpublished (2004)

COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 31 row: M column: 14.

FEATURES
source
1. .858
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea31M14"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/note="Vector: pCMV, SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by Agri. More glycerol clones held in -80."

ORIGIN
Query Match 9.3%; Score 228.6; DB 7; Length 858;
Best Local Similarity 66.5%; Pred. No. 1.2e-49;
Matches 393; Conservative 0; Mismatches 109; Indels 89; Gaps 1;

QY 652 GCGCTAGACGGTGGTTCCATCGCTTCCACAGGAGGTATTTGCTGTGATTGGAGAGAG 711
DB 1 GCGCGAGGCTGTAGTTCCATCTCTTCCAAATAGATTGCTGTATGTTGTCGGG 60

QY 712 TATTTGCAAAATGTGCAGAGAGCTTGTGTGAGAAGTTGCAGCTTATGCAGAAGCGGAA 771
DB 61 TATTTGCAAAATGTCCGAGAGATTAGTAGAGAACTGCACACCTTTTGCATTAGATGGCA 120

QY 772 GTGCTGTGAATATGGAAGCGAAGTTCTCTCAGATGACACTTCATGTCTATGCTGGTTGTCTC 831
DB 121 CTGCGGTGAACATGGAAGAAAAGTTTCTCACTGACTCTTGATGTTATCGGTCTCTCGG 180

QY 832 TTTTAACTACAAATTCGATCTTTTGACTACTGATAGTCTCTCATTTGAAGCTGTTTACA 891
DB 181 TATTTAATTAACCTTCGATTGCTGTGACACCGATAGCCCTGTCTATTGATGAGTTATA 240

QY 892 CTGCTCTTAAAGAGCTGAGCTTGGTTCTACTGATCTTCTGCCATATTCGAAGGCAAGTT 951
DB 241 CTGCTTTGAAAGAGCAGAAATACGATCCACGACATTTTACCTATTGGA----- 291

QY 952 TCTGTGTTTTTCTGTGTTGTTGATGTGTGGNACAAATTCGATTCTTGTATTTAG 1011
DB 292 ----- 291

QY 1012 AGGGTTTGGTTGTTTTTTCAGATCGATGTCATTTGTGAAGATAGTCCGAGACAGGTGAA 1071
DB 292 -----AGATTAGTCTTTTGTGCAAGNTAGTTCGAGACAATATAA 331

QY 1072 AGCTGAAAAGGCTGTAACTTTTGATAGGGAAACTGTTGAGACCTTTATTGCTAAGTGTAA 1131
DB 332 GGCCTGAAAAGCAGTTTACAGTCATTCGGAAAGCCGTTGGAAGAACTTATTGTAAGTGCAA 391

QY 1132 AGAATTTGCGAAGAGAGAGCGGAAGATCAATGATGAGGAGTATGTAATGATGCTGA 1191
DB 392 AGAGATTTGTGAAAAGAGGTTGAAAGGATCAACGAGGAGGAATATGTAATGATGCTGA 451

QY 1192 CCCAGTATCTCGCTTCTTCTGCTGCAAGCAGAGAAAGGTTTAAACTTT 1242
DB 452 TCCAAAGTATCCTTCGATTTCTTGCTGCAAGCAGGGAAGAGGTTTCAAGCTT 502

RESULT 15
CZ710414
LOCUS OC_Ba0031117.r OC_Ba Oryza coarctata genomic clone OC_Ba0031117
DEFINITION 3'---genomic survey sequence.
ACCESSION CZ710414
VERSION CZ710414
KEYWORDS GSS.
ORGANISM Oryza coarctata (Porteresia coarctata)
SOURCE Oryza coarctata
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Zingiberales; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 704)
AUTHORS Kim, H., Collura, K., Wisotski, M., Byrne, M., Stum, D., Smart, D.,
Rao, K., Luo, M., Jetty, R., Kudrna, D., Muller, C., Soderlund, C. and
Wing, R.
TITLE OMAP (Oryza Map Alignment Project)- Arizona Genomics Institute
JOURNAL Unpublished (2005)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0031 row: L column: 17
Seq primer: CAC TCA TTA GGC ACC CCA
Class: BAC ends.

FEATURES
Location/Qualifiers
1. .704
/organism="Oryza coarctata"
/mol_type="genomic DNA"
/db_xref="taxon:77588"
/clone="OC_Ba0031117"
/tissue_type="leaves"
/dev_stage="mature"
/lab_host="DH10B"
/clone_lib="OC_Ba"
/note="Vector: pAGIBAC1; Site 1: HindIII; Site 2: HindIII"

ORIGIN
Query Match 8.9%; Score 220.4; DB 10; Length 704;
Best Local Similarity 64.8%; Pred. No. 1.8e-47;
Matches 376; Conservative 0; Mismatches 171; Indels 33; Gaps 2;

QY 646 AGCGAGGCGTAGAGCGGTGGTTCCATCGCTTCCACAGGAGGTATTTGCTGTGATTGTGG 705
DB 124 AGGTGAGAGCTGCATCGTACCATCTCTACACAAACGATTTCTCTCAGTGATGGTTG 183

QY 706 AGAGATATTTGCAAAATGTGCAGAGAGGCTTGTGAGAAGTTGCAGCTTATGCAGAAG 765
DB 184 ACAGAGTTTTTTGTAATGTGCTGAGAGATTAGTGAGAAGCTTGAGACGCTCTGCTTTAA 243

QY 766 AGCGAAGTCTGTGATATGGAAGCGAAGTTCTCTCAGATGACACTTGTATGTCATTGGGT 825
DB 244 GTGGCAAACTGTAAATATGGAAGCAAGGTTCTCTCAAAAGACTTTTAGATGATTTGGTT 303

QY 826 TGCTCTTTTTTAACTACAAATTTTCGATTTCTTTGACTACTGATAGTCTCTGTTGAAGCTG 885
DB 304 TGTCTTGTTCATTTACAAATTTGATTTCCCTCAGCTCAGATAGTCCAGTTATTGATGCTG 363

QY 886 TTTACATGCTCTTTAAAGAGCTGAGCTTGGTTCTACTGATCTTCTGCGCATATTGAAGG 945
DB 364 TCTACATGCTCACTCAAGGAAGCAGAGCTTGGTTCTTACAGATCTTTTACCATATCGAAGG 423

QY 946 CAAGTTTCTGTGTTTTTTTCTGTGTTTGTGTTCA-----TTGTGTGGAAACAATTG 994
DB 424 TATATTCCTTTTGGCAATGTTTTTTTTTATATAAAAAAATCTCATTCGGTTGGCCTGTCTA 483


```

RESULT 9
US-11-052-554A-475
; Sequence 475, Application US/11052554A
; Publication No. US20050289866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD
; TITLE OF INVENTION: PROTEINS OF THERM
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,232
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763

```


QY 1243 TT 1244
DB 948 AT 949

RESULT 13

US-11-117-187-211

; Sequence 211, Application US/11117187

; Publication No. US20050266560A1

; GENERAL INFORMATION:

; APPLICANT: PREUSS, DAPHNE

; APPLICANT: COPENHAVER, GREGORY

; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS

; FILE REFERENCE: ARCD:309US

; CURRENT APPLICATION NUMBER: US/11/117,187

; PRIOR FILING DATE: 2005-04-28

; PRIOR APPLICATION NUMBER: US/09/531,120

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: 60/125,219

; PRIOR FILING DATE: 1999-03-18

; NUMBER OF SEQ ID NOS: 212

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 211

; LENGTH: 1082144

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-11-117-187-211

Query Match

Best Local Similarity 1.5%; Score 36.4; DB 7; Length 1082144;

Matches 70; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1145 AGAGAAGCGGAAGAATCAATGATGAGGAGTATGTAATGATGCTGACCCCAAGTATCCTG 1204

DB 525307 AGTGAATGAAATAATAATAAAGGCAATGAAGATGAAGTCAAACTTTTG 525366

QY 1205 GCTTCTGCTTGCAGCAGAGAAGGTTTAACTTTTCTTAAAGTTTATAAGCAAA 1264

DB 525367 ACTCTAATGTAATCTTATAGCAATAACCACTTTTCTACTGAAAGTTATAGAAAA 525426

QY 1265 TTTCGC 1270

DB 525427 TGAGGC 525432

RESULT 14

US-10-793-626-997

; Sequence 997, Application US/10793626

; Publication No. US20050255478A1

; GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PU3480US

; CURRENT APPLICATION NUMBER: US/10/793,626

; PRIOR FILING DATE: 2004-03-04

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 997

; LENGTH: 711

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: nucleic acid sequence

US-10-793-626-997

Query Match

Best Local Similarity 1.5%; Score 36.2; DB 6; Length 711;

Matches 101; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 2079 TCAAGTAACTCAGTAGAACACATCTTTTGACACAACTACTGAATCAAGATTAGTGGTT 2138
DB 484 TTAATAAATCTCTCCATTATAGCAATTTCAATGAGAATACTAAAGGATTTAAACAATTAATA 543
QY 2139 TTGATTAGGGAATTTAAAAGATGATTTCTTTTTCACAGATTTCATCCATTCACTGGA 2198
DB 544 GCATTAAGTGTATATTAGGAGAAATATCTGTAATCTAGGATTAATTATAGCTTTTAT 603
QY 2199 GGGCCTAGAAAAATGTTAGGGGATCAGTTTGCATTTGATGAGGCAATTTGGCACTCGCG 2258
DB 604 ATGAATATATCACCTGGTGGGCTCATTTGTTGTTACTATTGGTATTAAATGCTCATACG 663
QY 2259 GTGTTTCTTCAGCGGTTAAACGTTGAGCT 2287
DB 664 ATGATTATTTCAGAAAGTTAAAAAATTAAAGTT 692

RESULT 15

US-10-750-185-28428/c

; Sequence 28428, Application US/10750185

; Publication No. US20050260603A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen

; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 28428

; LENGTH: 1179

; TYPE: DNA

; ORGANISM: Bovine 19866881662207

US-10-750-185-28428

Query Match

Best Local Similarity 1.5%; Score 36.2; DB 6; Length 1179;

Matches 74; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 900 AAGAAGCTGAGCTTCGTTCTACTGATCTCTGCGCATATTCGAGGCAAGTTTCCTGTGT 959

DB 659 AAGGAATTTATTTGTAGGACTGATTCCTTCATCCATAATGTATACCATAGTTCTTTT 600

QY 960 TTTTCTGCTGTTTGTTCGTTGTTGGAACAAATTCGATTCCTTGTAAATGAGAGGTTTG 1019

DB 599 TTTTATCTGATTTTATTTGTTTTTTTTTTTAAATGAGGATAATGCTTTACAGTATTTT 540

QY 1020 GTTGTTTTTTCAGATC 1036

DB 539 GTTGTTCCTGCCAAAC 523

Search completed: January 13, 2006, 07:31:15

Job time : 689 secs

This Page Blank (uspto)